(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 November 2002 (14.11.2002)

PCT

(10) International Publication Number
WO 02/090566 A2

(51) International Patent Classification[7]: C12Q

(21) International Application Number: PCT/US02/13844

(22) International Filing Date: 3 May 2002 (03.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/288,564 3 May 2001 (03.05.2001) US

(71) Applicant: LEXIGEN PHARMACEUTICALS CORP. [US/US]; 125 Hartwell Avenue, Lexington, MA 02421 (US).

(72) Inventors: GILLIES, Stephen, D.; 159 Sunset Road, Carlisle, MA 01741 (US). LO, Kin-Ming; 6 Carol Lane, Lexington, MA 02420 (US). QIAN, Xiuqi; 122 Baker Avenue, Concord, MA 01742 (US).

(74) Agent: WALLER, Patrick, R., H.; Testa, Hurwitz & Thibeault, L.L.P., High Street Tower, 125 High Street, Boston, MA 02110 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/090566 A2

(54) Title: RECOMBINANT TUMOR SPECIFIC ANTIBODY AND USE THEREOF

(57) Abstract: The invention provides a family of antibodies that specifically bind the human epithelial cell adhesion molecule. The antibodies comprise modified variable regions, more specially, modified framework regions, which reduce their immunogenicity when administered to a human. The antibodies, when coupled to the appropriate moiety, may be used in the diagnosis, prognosis and treatment of cancer.

RECOMBINANT TUMOR SPECIFIC ANTIBODY AND USE THEREOF

RELATED APPLICATIONS

This application claims the benefit of and priority to U.S.S.N. 60/288,564, filed May 3, 2001, the disclosure of which is incorporated by reference herein.

FIELD OF THE INVENTION

The invention relates generally to recombinant antibodies. More particulary, the invention relates to recombinant antibodies that specifically bind human Epithelial Cell Adhesion Molecule, and to their use as diagnostic, prognostic and therapeutic agents.

BACKGROUND OF THE INVENTION

There has been significant progress in the development of antibody-based therapies over the years. For example, investigators have identified not only a variety of cancer-specific markers but also a variety of antibodies that bind specifically to those markers. Antibodies can be used to deliver certain molecules, for example, a toxin or an immune stimulatory moiety, for example, a cytokine, to a cancer cell expressing the marker so as to selectively kill the cancer cell (see, e.g., U.S. Patent Nos. 5,541,087; and 5,650,150).

The KS-1/4 antibody is a mouse-derived monoclonal antibody directed against human epithelial cell adhesion molecule (EpCAM). EpCAM is expressed at very low levels on the apical surface of certain epithelial cells. For example, EpCAM is expressed on intestinal cells on the cell surface facing toward ingested food and away from the circulation, where it would not be accessible to most proteins and cells of the immune system (Balzar *et al.* [1999] J. Mol. Med. 77:699-712).

Under certain circumstances, however, EpCAM is highly expressed on certain cells, for example, tumor cells of epithelial origin. Typically, these tumor cells have lose their polarity with the result that EpCAM is expressed over the entire surface of the cell.

Thus, EpCAM is a convenient tumor-specific marker for directing antibody-based immune-stimulatory moieties to tumor cells (Simon *et al.* [1990] Proc. Nat. Acad. Sci. USA 78:2755-2759; Perez *et al.* [1989] J Immunol. 142:3662-3667).

However, antibodies can have an associated immunogenicity in the host mammal. This is more likely to occur when the antibodies are not autologous. Consequently, the effectiveness of antibody-based therapies often is by an immunogenic response directed against the antibody. The immunogenic response typically is increased when the antibody is derived in whole or in part from a mammal different than the host mammal, e.g., when the antibody is derived from a mouse and the recipient is a human. Accordingly, it may be helpful to modify mouse-derived antibodies to more closely resemble human antibodies, so as to reduce or minimize the immunogenicity of the mouse-derived antibody.

Although a variety of approaches have been developed, including, for example, chimeric antibodies, antibody humanization and antibody veneering,Accordingly, there is a need in the art for antibodies that bind to cancer specific markers and that have reduced immunogenicity when administered to a human. Further, there is a need in the art for antibodies that deliver toxins or immune stimulatory moieties, for example, as fusion proteins or immune conjugates to a cancer specific marker to selectively kill the tumor cell.

SUMMARY OF THE INVENTION

The present invention is based, in part, upon the identification of recombinant antibodies that specifically bind human EpCAM but are less immunogenic in humans than the template, murine anti-EpCAM antibodies. In particular, the invention provides recombinant KS antibodies in which the amino acid sequences defining one or more framework regions and/or complementarity determining regions have been modified to reduce their immunogenicity in humans.

As used herein, the terms "antibody" and "immunoglobulin" are understood to mean (i) an intact antibody (for example, a monoclonal antibody or polyclonal antibody), (ii) antigen binding portions thereof, including, for example, an Fab fragment, an Fab'

fragment, an $(Fab')_2$ fragment, an Fv fragment, a single chain antibody binding site, an sFv, (iii) bi-specific antibodies and antigen binding portions thereof, and (iv) multi-specific antibodies and antigen binding portions thereof.

As used herein, the terms "bind specifically," "specifically bind" and "specific binding" are understood to mean that the antibody has a binding affinity for a particular antigen of at least about 10^6 M^{-1}, more preferably, at least about 10^7 M^{-1}, more preferably at least about 10^8 M^{-1}, and most preferably at least about 10^{10} M^{-1}.

As used herein, the terms "Complementarity-Determining Regions" and "CDRs" are understood to mean the hypervariable regions or loops of an immunoglobulin variable region that interact primarily with an antigen. The immunoglobulin heavy chain variable region (V_H) and immunoglobulin light chain variable region (V_L) both contain three CDRs interposed between framework regions, as shown in Figure 1. For example, with reference to the amino acid sequence defining the immunoglobulin light chain variable of the of the KS-1/4 antibody as shown in SEQ ID NO: 1, the CDRs are defined by the amino acid sequences from Ser24 to Leu33 (CDR1), from Asp49 to Ser55 (CDR2), and from His88 to Thr96 (CDR3). With reference to the amino acid sequence defining the immunoglobulin heavy chain variable region of the KS-1/4 antibody as shown in SEQ ID NO: 2, the CDRs are defined by the amino acid sequences from Gly26 to Asn35 (CDR1), from Trp50 to Gly66 (CDR2), and from Phe99 to Tyr105 (CDR3). The corresponding CDRs of the other antibodies described herein are shown in Figures 1A-1C after alignment with the corresponding KS-1/4 heavy or light chain sequence.

As used herein, the terms "Framework Regions" and "FRs" are understood to mean the regions an immunoglobulin variable region adjacent to the Complementarity-Determining Regions. The immunoglobulin heavy chain variable region (V_H) and immunoglobulin light chain variable region (V_L) both contain four FRs, as shown in Figure 1. For example, with reference to the amino acid sequence defining the immunoglobulin light chain variable of the of the KS-1/4 antibody as shown in SEQ ID NO: 1, the FRs are defined by the amino acid sequences from Gln1 to Cys23 (FR1), from Trp34 to Phe 48 (FR2), from Gly56 to Cys87 (FR3), and from Phe97 to Lys106 (FR4).

With reference to the amino acid sequence defining the immunoglobulin heavy chain variable region of the KS-1/4 antibody as shown in SEQ ID NO: 2, the FRs are defined by the amino acid sequences from Gln1 to Ser25 (FR1), from Trp36 to Gly49 (FR2), from Arg67 to Arg98 (FR3), and from Trp106 to Ser116 (FR4). The FRs of the other antibodies described herein are shown in Figures X and Y after alignment with the corresponding KS-1/4 heavy or light chain sequence.

As used herein, the term "KS antibody" is understood to mean an antibody that binds specifically to the same human EpCAM antigen bound by murine antibody KS-1/4 expressed by a hybridoma (see, for example, Cancer Res. 1984, 44 ((2):681-7). The KS antibody preferably comprises (i) an amino acid sequence of SASSSVSY (amino acids 24-31 of SEQ ID NO: 1) defining at least a portion of an immunoglobulin light chain CDR1 sequence, (ii) an amino acid sequence of DTSNLAS (amino acids 49-55 of SEQ ID NO: 1) defining at least a portion of an immunoglobulin light chain CDR2 sequence, (iii) an amino acid sequence of HQRSGYPYT (amino acids 88-96 of SEQ ID NO: 1) defining at least a portion of an immunoglobulin light chain CDR3 sequence, (iv) an amino acid sequence of GYTFTNYGMN (amino acids 26-35 of SEQ ID NO: 2) defining at least a portion of an immunoglobulin heavy chain CDR1 sequence, (v) an amino acid sequence of WINTYTGEPTYAD (amino acids 50-62 of SEQ ID NO: 2) defining at least a portion of an immunoglobulin heavy chain CDR2 sequence, or (vi) an amino acid sequence of SKGDY (amino acids 101-105 of SEQ ID NO: 2) defining at least a portion of an immunoglobulin heavy chain CDR3 sequence, or any combination of the foregoing.

In one aspect, the invention provides a recombinant antibody that specifically binds EpCAM, wherein the antibody comprises an amino acid sequence, a portion of which defines a framework region in an immunoglobulin V_L domain. In one embodiment, the framework region (FR1) is defined by amino acid residues 1-23 of SEQ ID NO: 5, wherein Xaa1 is Q or E, Xaa3 is L or V, Xaa10 is I or T, Xaa11 is M or L, Xaa13 is A or L, Xaa18 is K or R, or Xaa21 is M or L, provided that at least one of the amino acid residues at positions Xaa1, Xaa3, Xaa10, Xaa11, Xaa13, Xaa18, or Xaa21 is not the same as the amino acid at the corresponding position in SEQ ID NO: 1. The amino acids at each of the positions are denoted by the standard single letter code.

In another embodiment, the framework region (FR2) is defined by amino acid residues 34-48 of SEQ ID NO: 5, wherein Xaa41 is S or Q, Xaa42 is S or A, Xaa45 is P or L, or Xaa46 is W or L, provided that at least one of the amino acid residues at positions Xaa41, Xaa42, Xaa45, or Xaa46 is not the same as the amino acid at the corresponding position in SEQ ID NO: 1.

In another embodiment, the framework region (FR3) is defined by amino acid residues 56-87 of SEQ ID NO: 5, wherein Xaa57 is F or I, Xaa69 is S or D, Xaa71 is S or T, Xaa73 is I or T, Xaa77 is M or L, Xaa79 is A or P, Xaa82 is A or F, or Xaa84 is T or V, provided that at least one of the amino acid residues at positions Xaa57, Xaa69, Xaa71, Xaa73, Xaa77, Xaa79, Xaa82, or Xaa84 is not the same as the amino acid at the corresponding position in SEQ ID NO: 1.

In another aspect, the invention provides a recombinant antibody that specifically binds EpCAM, wherein the antibody comprises an amino acid sequence, a portion of which defines a framework region in an immunoglobulin V_L domain. In one embodiment, the framework region (FR1) is defined by amino acid residues 1-25 of SEQ ID NO: 6, wherein Xaa2 is I or V, Xaa9 is P or A, Xaa11 is L or V, or Xaa17 is T or S, provided that at least one of the amino acid residues at positions Xaa2, Xaa9, Xaa11 or Xaa17 is not the same as the amino acid at the corresponding position in SEQ ID NO: 2.

In another embodiment, the framework region (FR2) is defined by amino acid residues 36-49 of SEQ ID NO: 6, wherein Xaa38 is K or R, Xaa40 is T or A, or Xaa46 is K or E, provided that at least one of the amino acid residues at positions Xaa38, Xaa40, Xaa46 is not the same as the amino acid at the corresponding position in SEQ ID NO: 2.

In another embodiment, the framework region (FR3) is defined by amino acid residues 67-98 of SEQ ID NO: 6, wherein Xaa68 is F or V, Xaa69 is A or T, Xaa70 is F or I, Xaa73 is E or D, Xaa76 is A or T, Xaa80 is F or Y, Xaa83 is I or L, Xaa84 is N or S, Xaa85 is N or S, Xaa88 is N, A or S, Xaa91 is M or T, or Xaa93 is T or V, provided that at least one of the amino acid residues at positions Xaa68, Xaa69, Xaa70, Xaa73, Xaa76, Xaa80, Xaa83, Xaa84, Xaa85, Xaa88, Xaa91 or Xaa93 is not the same as the amino acid at the corresponding position in SEQ ID NO: 2. In another embodiment, the framework

region (FR4) is defined by amino acid residues 106-116 of SEQ ID NO: 6, wherein Xaa108 is Q or T.

In another embodiment, the immunoglobulin V_L domain comprises an FR1 sequence selected from the group consisting of: (i) amino acid residues 1-23 of SEQ ID NO: 9; and (ii) amino acid residues 1-23 of SEQ ID NO: 8. In another embodiment, the immonoglobulin V_H domains comprises an FR sequence defined by amino acid residues 1-25 of SEQ ID NO: 18 and or an FR sequence defined by amino acid residues 67-98 of SEQ ID NO: 18. More preferably, the V_L domain comprises an amino acid sequence defined by amino acids 1-106 of SEQ ID NO: 9 and/or the V_H domain comprises an amino acid sequence defined by amino acids 1-116 of SEQ ID NO: 18.

Furthermore, the antibody optionally may include an amino acid sequence defining at least a portion of a CDR sequence including, for example, (i) amino acid residues 24-31 of SEQ ID NO: 1; (ii) amino acid residues 49-55 of SEQ ID NO: 1; and/or (iii) amino acid residues 88-96 of SEQ ID NO: 1. Similarly, the antibody optionally may include an amino acid sequence defining at least a portion of a CDR sequence including, for example, (i) amino acid residues 26-35 of SEQ ID NO: 2; (ii) amino acid residues 50-62 of SEQ ID NO: 2; and/or iii) amino acid residues 101-105 of SEQ ID NO: 2.

In another embodiment, the antibody comprises the antigen targeting portion of an antibody-cytokine fusion protein. The cytokine preferably is an interleukin and more preferably is interleukin-2.

In another aspect, the invention provides an expression vector encoding at least a portion of the antibody of the invention. In a preferred embodiment, the expression vector comprises the nucleotide sequence set forth in SEQ ID NO: 40.

In another aspect, the invention provides a method of diagnosing, prognosing and/or treating a human patient having a disease associated with over-expression of EpCAM (for example, a disease in which EpCAM is present at a higher level in diseased tissue relative to tissue without that disease). The method comprises administering one of

the antibodies of the invention to an individual in need of such diagnosis, prognosis or treatment.

The antibody optionally includes a diagnostic and/or therapeutic agent attached thereto. The agent may be fused to the antibody to produce a fusion protein. Alternatively, the agent may be chemically coupled to the antibody to produce an immuno-conjugate. It is contemplated that the agent may include, for example, a toxin, radiolabel, cytokine, imaging agent or the like. In a preferred embodiment, the antibody of the invention is fused as a fusion protein to a cytokine. Preferred cytokines preferably include interleukins such as interleukin-2 (IL-2), IL-4, IL-5, IL-6, IL-7, IL-10, IL-12, IL-13, IL-14, IL-15, IL-16 and IL-18, hematopoietic factors such as granulocyte-macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF) and erythropoeitin, tumor necrosis factors (TNF) such as TNFα, lymphokines such as lymphotoxin, regulators of metabolic processes such as leptin, interferons such as interferon α, interferon β, and interferon γ, and chemokines. Preferably, the antibody-cytokine fusion protein displays cytokine biological activity.

DESCRIPTION OF THE DRAWINGS

Figures 1A, 1B and 1C show an alignment of light and heavy chain variants and consensus sequences of KS antibodies. The immunoglobulin Framework Regions (FR1-FR4) are denoted by -. The immunoglobulin Complementarity Determining Regions (CDR1-CDR3) are denoted by *. Individual KS antibody light chain V region segments are referred to as "VK," wherein K refers to the fact that the light chain is a kappa chain. Individual KS antibody heavy chain V region segments are referred to as "V_H." Substitutable amino acids are denoted by "X" in the consensus sequences.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides recombinant antibodies that specifically bind human Epithelial Cell Adhesion Molecule (EpCAM). Preferred antibodies of the invention have altered variable regions that result in reduced immunogenicity in humans.

Antibody variable regions of the invention are particularly useful to target antibodies and antibody fusion proteins to tumor tissues that over-express EpCAM in human patients. In preferred embodiments, an antibody of the invention is fused to a cytokine to produce an immuno-cytokine.

Protein sequences of the invention

The present invention discloses a family of antibody variable region or V region sequences that, when appropriately heterodimerized, bind to human epithelial cell adhesion molecule (EpCAM) also known as KS antigen or KSA. Preferred proteins of the invention are useful for treating human patients as described herein. Accordingly, preferred KS antibody variants are humanized, deimmunized, or both, in order to reduce their immunogenicity when administered to a human. According to the invention, murine KS antibodies can be deimmunized or humanized, for example, by using deimmunization methods in which potential T cell epitopes are eliminated or weakened by introduction of mutations that reduce binding of a peptide epitope to an MHC Class II molecule (see, for example WO98/52976, and WO00/34317), or by using methods in which non-human T cell epitopes are mutated so that they correspond to human self epitopes that are present in human antibodies (see, for example, U.S. Patent No. 5,712,120).

I. Variable Light Chain

The recombinant anti-EpCAM antibody has an immunoglobulin variable light chain sequence having the following amino acid sequence:

X-I-X-L-T-Q-S-P-A-X-X-X-X-S-P-G-X-X-X-T-X-T-C- S-A-S-S-S-V-S-T-X-L-W-Y-X-Q-K-P-G-X-X-P-K-X-X-I-X-D-T-S-N-L-A-S-G-X-P-X-R-F-S-G-S-G-S-G-T-X-Y-X-L-X-I-X-S-X-E-X-E-D-X-A-X-Y-Y-C-H-Q-R-S-G-Y-P-Y-T-F-G-G-G-T-K-X-E-I-K (SEQ ID NO: 3).

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light chain FR1, which is represented by residues 1 to 23 of SEQ ID NO: 3, namely, X-I-X-L-T-Q-S-P-A-X-X-X-X-S-P-G-X-X-

X-T-X-T-C. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR1 region: Q or E at position Xaa1; L or V at position Xaa3; I, T or S at position Xaa10; M or L at position Xaa11; S or A at position Xaa12; A, L or V at position Xaa13; E or Q at position Xaa17, K or R at position Xaa18, V or A at position Xaa19; and, M, L or I at position Xaa21. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR1 region: E at position Xaa1; V at position Xaa3; T or S at position Xaa10; L at position Xaa11; A at position Xaa12; L or V at position Xaa13; Q at position Xaa17, R at position Xaa18, A at position Xaa19; and, L or I at position Xaa21.

In another embodiment, the recombinant anti-EpCAM antibody of the invention has an amino acid sequence defining an immunoglobulin light chain CDR1, which is represented by residues 24 to 33 of SEQ ID NO: 3, namely S-A-S-S-S-V-S-T-X-L. More particularly, the recombinant anti-EpCAM antibody of the invention has one of the following amino acids in the CDR1 region: M or I at position Xaa32. More preferably, the recombinant anti-EpCAM antibody has an amino acid substitution in the CDR1 region, for example, I at position Xaa32.

In another embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light chain FR2, which is represented by residues 34 to 48 of SEQ ID NO: 3, namely W-Y-X-Q-K-P-G-X-X-P-K-X-X-I-X. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR2 region: Q or L at position Xaa36; S or Q at position Xaa41; S, A or P at position Xaa42; P or L at position Xaa45; W or L at position Xaa46; and, F or Y at position Xaa48. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR2 region: L at position Xaa36; Q at position Xaa41; A or P at position Xaa42; L at position Xaa45; L at position Xaa46; and, Y at position Xaa48.

In another embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light chain FR3, which is represented by residues 56 to 87 of SEQ ID NO: 3, namely, G-X-P-X-R-F-S-G-S-G-S-G-T-X-Y-X-L-X-I-X-S-X-E-X-E-D-X-A-X-Y-Y-C. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR3 region: F or I at position Xaa57;

A or S at position Xaa59; S, D or T at position Xaa69; I or T at position Xaa71; I or T at position Xaa73; S or N at position Xaa75; M or L at position Xaa77; A or P at position Xaa79; A or F at position Xaa82; and, T or V at position Xaa84. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitution in the FR3 region: I at position Xaa57; S at position Xaa59; D or T at position Xaa69; T at position Xaa71; T at position Xaa73; N at position Xaa75; L at position Xaa77; P at position Xaa79; F at position Xaa82; and, V at position Xaa84.

In another embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light chain FR4, which is represented by residues 97 to 106 of SEQ ID NO: 3, namely, F-G-G-G-T-K-X-E-I-K. More particularly, the recombinant anti-EpCAM antibody of the invention has at least one of the following amino acids in the FR4 region, for example, L or V at position Xaa103. Accordingly, the recombinant anti-EpCAM antibody of the invention has an amino acid substitution in the FR4 region, for example, V at position Xaa103.

II. Variable Heavy Chain

The recombinant anti-EpCAM antibody has an immunoglobulin variable heavy chain sequence having the following amino acid sequence:

Q-X-Q-L-V-Q-S-G-X-E-X-K-K-P-G-X-X-V-K-I-S-C-K-A-S-G-Y-T-F-T-N-Y-G-M-N-W-V-X-Q-X-P-G-X-G-L-X-W-M-G-W-I-N-T-Y-T-G-E-P-T-Y-A-D-X-F-X-G-R-X-X-X-X-X-X-T-S-X-S-T-X-X-L-Q-X-X-X-L-R-X-E-D-X-A-X-Y-F-C-V-R-F-X-S-K-G-D-Y-W-G-X-G-T-X-V-T-V-S-S (SEQ ID NO: 4)

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain FR1, which is represented by residues 1 to 25 of SEQ ID NO: 4, namely Q-X-Q-L-V-Q-S-G-X-E-X-K-K-P-G-X-X-V-K-I-S-C-K-A-S. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR1 region: I or V at position Xaa2; P or A at position Xaa9; L or V at position Xaa11; E or S at position Xaa16; and, T or S at position Xaa17. More preferably, the recombinant anti-EpCAM antibody has at least one of the

following amino acid substitutions in the FR1 region: V at position Xaa2; A at position Xaa9; V at position Xaa11; S at position Xaa16; and, S at position Xaa17.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain FR2, which is represented by residues 36 to 49 of SEQ ID NO: 4, W-V-X-Q-X-P-G-X-G-L-X-W-M-G. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR2 region: K or R at position Xaa38; T or A at position Xaa40; K or Q at position Xaa43; and, K or E at position Xaa46. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR2 region: R at position Xaa38; A at position Xaa40; Q at position Xaa43; and, E at position Xaa46.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain CDR2, which is represented by residues 50 to 66 of SEQ ID NO: 4, namely W-I-N-T-Y-T-G-E-P-T-Y-A-D-X-F-X-G. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the CDR2 region: D or K at position Xaa63; and, K or Q at position Xaa65. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the CDR2 region: K at position Xaa63; and, Q at position Xaa65.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain FR3, which is represented by residues 67 to 98 of SEQ ID NO: 4, namely R-X-X-X-X-X-X-T-S-X-S-T-X-X-L-Q-X-X-X-L-R-X-E-D-X-A-X-Y-F-C-V-R. More particularly, the recombinant anti-EpCAM antibody of the invention has at least one of the following amino acids in the FR3 region: F or V at position Xaa68, A, T or V at position Xaa69; F or I at position Xaa70; S or T at position Xaa71; L or A at position Xaa72; E or D at position Xaa73; A or T at position Xaa76; A or L at position Xaa79; F or Y at position Xaa80; I or L at position Xaa83; N or S at position Xaa84; N or S at position Xaa85; N, A or S at position Xaa88; M or T at position Xaa91; and, T or V at position Xaa93. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR3 region: V at position Xaa68, T or V at position Xaa69; I at position Xaa70; T at position

Xaa71; A at position Xaa72; D at position Xaa73; T at position Xaa76; L at position Xaa79; Y at position Xaa80; L at position Xaa83; S at position Xaa84; S at position Xaa85; A or S at position Xaa88; T at position Xaa91; and, V at position Xaa93.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain CDR3, which is represented by residues 99 to 105 of SEQ ID NO: 4, namely F-X-S-K-G-D-Y. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the CDR3 region, for example, I or M at position Xaa100. More preferably, the recombinant anti-EpCAM antibody has an amino acid substitution in the CDR3 region, for example, M at position Xaa100.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain FR4, which is represented by residues 106 to 116 of SEQ ID NO: 4, namely W-G-X-G-T-X-V-T-V-S-S. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR4 region: Q or T at position Xaa108; and, S or T at position X111. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR4 region: T at position Xaa108; and, T at position X111.

III. Refined Variable Light Chain

In another embodiment, the recombinant anti-EpCAM antibody has an immunoglobulin variable light chain sequence having the following amino acid sequence:

X-I-X-L-T-Q-S-P-A-X-X-S-X-S-P-G-E-X-V-T-X-T-C-S-A-S-S-S-V-S-Y-M-L-W-Y-Q-Q-K-P-G-X-X-P-K-X-X-I-F-D-T-S-N-L-A-S-G-X-P-A-R-F-S-G-S-G-S-G-T-X-Y-X-L-X-I-S-S-X-E-X-E-D-X-A-X-Y-Y-C -H-Q-R-S-G-Y-P-Y-T-F-G-G-G-T-K-L-E-I-K (SEQ ID NO: 5)

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light chain FR1, which is represented by

residues 1 to 23 of SEQ ID NO: 5, namely X-I-X-L-T-Q-S-P-A-X-X-S-X-S-P-G-E-X-V-T-X-T-C. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR1 region: Q or E at position Xaa1; L or V at position Xaa3; I or T at position Xaa10; M or L at position Xaa11; A or L at position Xaa13; K or R at position Xaa18; and, M or L at position Xaa21. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR1 region: E at position Xaa1; V at position Xaa3; T at position Xaa10; L at position Xaa11; L at position Xaa13; R at position Xaa18; and, L at position Xaa21.

In another preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light FR1 having at least one of the following amino acids in the FR1 region: Q or E at position Xaa1; A or L at position Xaa11; and, M or L at position Xaa21. More preferably, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light FR1 having at least one of the following substitutions in the FR1 region: E at position Xaa1; L at position Xaa11; and, L at position Xaa21.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light chain FR2, which is represented by residues 34 to 48 of SEQ ID NO: 5, namely W-Y-Q-Q-K-P-G-X-X-P-K-X-X-I-F. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR2 region: S or Q at position Xaa41; S or A at position Xaa42; P or L at position Xaa45; and, W or L at position Xaa46. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR2 region: Q at position Xaa41; A at position Xaa42; L at position Xaa45; and, L at position Xaa46.

In another preferred embodiment, the recombinant anti-EpCAM antibody of the invention has an amino acid sequence defining an immunoglobulin light FR2 having at least one of the following amino acids in the FR2 region: S or A at position Xaa42; P or L at position Xaa45; and, W or L at position Xaa46. More preferably, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light FR2 having at least one of the following substitutions in the FR2 region: A at position Xaa42; L at position Xaa45; and, L at position Xaa46.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light chain FR3, which is represented by residues 56 to 87 of SEQ ID NO: 5, namely G-X-P-A-R-F-S-G-S-G-S-G-T-X-Y-X-L-X-I-S-S-X-E-X-E-D-X-A-X-Y-Y-C. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR3 region: F or I at position Xaa57; S or D at position Xaa69; S or T at position Xaa71; I or T at position Xaa73; M or L at position Xaa77; A or P at position Xaa79; A or F at position Xaa82; and, T or V at position Xaa84. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitution in the FR3 region: I at position Xaa57; D at position Xaa69; T at position Xaa71; T at position Xaa73; L at position Xaa77; P at position Xaa79; F at position Xaa82; and, V at position Xaa84.

In another preferred embodiment, the recombinant anti-EpCAM antibody of the invention has an amino acid sequence defining an immunoglobulin light FR3 having at least one of the following amino acids in the FR3 region: F or I at position Xaa57; S or D at position Xaa69; A or P at position Xaa79; A or F at position Xaa82; and, T or V at position Xaa84. More preferably, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin light FR3 having at least one of the following substitutions in the FR3 region: I at position Xaa57; D at position Xaa69; P at position Xaa79; F at position Xaa82; and, V at position Xaa84.

IV. Refined Variable Heavy Chain

The recombinant anti-EpCAM antibody has an immunoglobulin variable heavy chain sequence having the following amino acid sequence:

Q-X-Q-L-V-Q-S-G-X-E-X-K-K-P-G-E-X-V-K-I-S-C-K-A-S-G-Y-T-F-T-N-Y-G-M-N-W-V-X-Q-X-P-G-K-G-L-X-W-M-G- W-I-N-T-Y-T-G-E-P-T-Y-A-D-X-F-X-G-R-X-X-X-S-L-X-T-S-X-S-T-A-X-L-Q-X-X-X-L-R-X-E-D-X-A-X-Y-F-C-V-R-F-I-S-K-G-D-Y-W-G-Q-G-T-S-V-T-V-S-S (SEQ ID NO: 6)

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain FR1, which is represented by

residues 1 to 25 of SEQ ID NO: 6, namely Q-X-Q-L-V-Q-S-G-X-E-X-K-K-P-G-E-X-V-K-I-S-C-K-A-S. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR1 region: I or V at position Xaa2; P or A at position Xaa9; L or V at position Xaa11; and, T or S at position Xaa17. Accordingly, a recombinant anti-EpCAM antibody of the invention has at least one of the following amino acid substitution in the FR1 region: V at position Xaa2; A at position Xaa9; V at position Xaa11; and, S at position Xaa17.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy FR1 having at least one of the following amino acids in the FR1 region: I or V at position Xaa2; P or A at position Xaa9; and, L or V at position Xaa11. Accordingly, a recombinant anti-EpCAM antibody of the invention has an amino acid sequence defining an immunoglobulin heavy FR1 having at least one of the following substitutions in the FR1 region: V at position Xaa2; A at position Xaa9; and, V at position Xaa11.

In another embodiment, a recombinant anti-EpCAM antibody of the invention has an amino acid sequence defining an immunoglobulin heavy chain FR2, which is represented by residues 36 to 49 of SEQ ID NO: 6, namely W-V-X-Q-X-P-G-K-G-L-X-W-M-G. More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitution in the FR2 region: K or R at position Xaa38; T or A at position Xaa40; and, K or E at position Xaa46. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitution in the FR2 region: R at position Xaa38; A at position Xaa40; and, E at position Xaa46.

In another preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy FR2 having the following amino acids in the FR1 region, for example, K or E at position Xaa46. More preferably, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy FR2 having an amino acid substitution in the FR1 region, for example, E at position Xaa46.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain CDR2, which is represented by residues 50 to 66 of SEQ ID NO: 6, namely W-I-N-T-Y-T-G-E-P-T-Y-A-D-X-F-X-G.

More particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the CDR2 region: D or K at position Xaa63; and, K or Q at position Xaa65. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the CDR2 region: K at position Xaa63; and, Q at position Xaa65.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain FR3, which is represented by residues 67 to 98 of SEQ ID NO: 6, namely R-X-X-X-S-L-X-T-S-X-S-T-A-X-L-Q-X-X-X-L-R-X-E-D-X-A-X-Y-F-C-V-R. More particularly, the recombinant anti-EpCAM antibody of the invention has at least one of the following amino acids in the FR3 region: F or V at position Xaa68; A or T at position Xaa69; F or I at position Xaa70; E or D at position Xaa73; A or T at position Xaa76; F or Y at position Xaa80; I or L at position Xaa83; N or S at position Xaa84; N or S at position Xaa85; N, A or S at position Xaa88; M or T at position Xaa91; and, T or V at position Xaa93. More preferably, the recombinant anti-EpCAM antibody has at least one of the following amino acid substitutions in the FR3 region: V at position Xaa68; T at position Xaa69; I at position Xaa70; D at position Xaa73; T at position Xaa76; Y at position Xaa80; L at position Xaa83; S at position Xaa84; S at position Xaa85; A or S at position Xaa88; T at position Xaa91; and, V at position Xaa93.

In another preferred embodiment, the recombinant anti-EpCAM antibody of the invention has an amino acid sequence defining an immunoglobulin heavy chain FR3 having at least one of the following amino acids in the FR3 region: F or V at position Xaa68; E or D at position Xaa73; N or S at position Xaa84; N or S at position Xaa85; N or A at position Xaa88; and, T or V at position Xaa93. More preferrably, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy FR3 having at least one of the following substitutions in the FR3 region: V at position Xaa68; D at position Xaa73; S at position Xaa84; S at position Xaa85; A at position Xaa88; and, V at position Xaa93.

In a preferred embodiment, the recombinant anti-EpCAM antibody has an amino acid sequence defining an immunoglobulin heavy chain FR4, which is represented by residues 106 to 116 of SEQ ID NO: 6, namely W-G-X-G-T-S-V-T-V-S-S. More

particularly, the recombinant anti-EpCAM antibody has at least one of the following amino acids in the FR4 region, for example, Q or T at position Xaa108. More preferably, the recombinant anti-EpCAM antibody has an amino acid substitution in the FR4 region, for example, T at position Xaa108.

Accordingly, preferred V regions contain substitutions in FR domains of V_H and/or VK regions corresponding to murine KS-1/4 variable regions. In addition, preferred V regions of the invention do not include insertions or deletions of amino acids relative to the murine KS-1/4 variable regions.

Preferred variants include proteins having variable regions with greater than 80% identity/homology murine KS-1/4. The amino acid sequence of murine KS variable region or a portion thereof may be used as a reference sequence to determine whether a candidate sequence possesses sufficient amino acid similarity to have a reasonable expectation of success in the methods of the present invention. Preferably, variant sequences are at least 70% similar or 60% identical, more preferably at least 75% similar or 65% identical, and most preferably 80% similar or 70% identical to a murine KS variable heavy or light chain FR or CDR.

To determine whether a candidate peptide region has the requisite percentage similarity or identity to a murine KS sequence, the candidate amino acid sequence and murine KS sequence are first aligned using the dynamic programming algorithm described in Smith and Waterman (1981) J. Mol. Biol. 147:195-197, in combination with the BLOSUM62 substitution matrix described in Figure 2 of Henikoff and Henikoff (1992) PNAS 89:10915-10919. For the present invention, an appropriate value for the gap insertion penalty is -12, and an appropriate value for the gap extension penalty is -4. Computer programs performing alignments using the algorithm of Smith-Waterman and the BLOSUM62 matrix, such as the GCG program suite (Oxford Molecular Group, Oxford, England), are commercially available and widely used by those skilled in the art. Once the alignment between the candidate and reference sequence is made, a percent similarity score may be calculated. The individual amino acids of each sequence are compared sequentially according to their similarity to each other. If the value in the BLOSUM62 matrix corresponding to the two aligned amino acids is zero or a negative number, the pairwise similarity score is zero; otherwise the pairwise similarity score is

1.0. The raw similarity score is the sum of the pairwise similarity scores of the aligned amino acids. The raw score is then normalized by dividing it by the number of amino acids in the smaller of the candidate or reference sequences. The normalized raw score is the percent similarity. Alternatively, to calculate a percent identity, the aligned amino acids of each sequence are again compared sequentially. If the amino acids are non-identical, the pairwise identity score is zero; otherwise the pairwise identity score is 1.0. The raw identity score is the sum of the identical aligned amino acids. The raw score is then normalized by dividing it by the number of amino acids in the smaller of the candidate or reference sequences. The normalized raw score is the percent identity. Insertions and deletions are ignored for the purposes of calculating percent similarity and identity. Accordingly, gap penalties are not used in this calculation, although they are used in the initial alignment.

The invention also discloses methods for assaying the expression of KS antibodies from cells such as mammalian cells, insect cells, plant cells, yeast cells, other eukaryotic cells or prokaryotic cells (see Example 1). In a preferred method, KS antibody V regions are expressed as components of an intact human antibody, and the expression of the antibody from a eukaryotic cell line assayed by an ELISA that detects the human Fc region. To precisely quantify binding of a KS antibody to EpCAM, a Biacore assay may be used.

Treatment of human disease with KS antibody fusion proteins

The invention also discloses the sequences of KS antibody-IL2 fusion proteins that are useful in treating human disease, such as cancer. Certain KS antibody-IL2 fusion proteins, such as KS-1/4-IL2 (see, for example, Construct 3 in Example X), may be used to treat human patients with cancer, with surprisingly little immune response against the antibody.

It is found that, during treatment of human cancers with KS-1/4(VH2/VK1)-IL2, even less immunogenicity is seen than with KS-1/4(Construct 3)-IL2. Specifically, during a clinical trial, patients with anti-idiotypic antibodies and antibody directed against the antibody-IL2 junction or against the IL-2 moiety are seen at an even lower frequency

than with KS-1/4(Construct 3)-IL2. Antibody variable regions of the invention can also be fused to other cytokines, for example, interleukins 1, 2, 6, 10, or 12; interferons alpha and beta; TNF, and INF gamma. The invention may be more fully understood by reference to the following non-limiting examples

EXAMPLES

Example 1. Methods and reagents for expressing KS antibodies and assaying their antigen-binding activity

1A. Cell culture and transfection

The following general techniques were used in the subsequent Examples. For transient transfection, plasmid DNA was introduced into human kidney 293 cells by co-precipitation of plasmid DNA with calcium phosphate [Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY].

In order to obtain stably transfected clones, plasmid DNA was introduced into the mouse myeloma NS/0 cells by electroporation. NS/0 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. About 5×10^6 cells were washed once with PBS and resuspended in 0.5 ml phosphate buffer solution (PBS). Ten μg of linearized plasmid DNA was then incubated with the cells in a Gene Pulser Cuvette (0.4 cm electrode gap, BioRad) for 10 minutes on ice. Electroporation was performed using a Gene Pulser (BioRad) with settings at 0.25 V and 500 μF. Cells were allowed to recover for 10 minutes on ice, after which they were resuspended in growth medium and then plated onto two 96-well plates. Stably transfected clones were selected by growth in the presence of 100 nM methotrexate (MTX), which was introduced two days post-transfection. The cells were fed every 3 days for two to three more times, and MTX-resistant clones appeared in 2 to 3 weeks. Supernatants from clones were assayed by anti-human Fc ELISA to identify high producers [Gillies *et al.* (1989) J. Immunol. Methods 125:191]. High producing clones were isolated and propagated in growth medium containing 100 nM MTX.

1B. ELISAs

Three different ELISAs were used to determine the concentrations of protein products in the supernatants of MTX-resistant clones and other test samples. The anti-huFc ELISA was used to measure the amount of human Fc-containing proteins, e.g., chimeric antibodies. The anti-hu kappa ELISA was used to measure the amount of kappa light chain (of chimeric or human immunoglobulins). The anti-muFc ELISA was used to measure the amount of muFc-containing proteins in test samples (see Example 1C below).

The anti-huFc ELISA is described in detail below.

A. Coating plates

ELISA plates were coated with AffiniPure goat anti-human IgG (H+L) (Jackson Immuno Research) at 5 μg/ml in PBS, and 100 μl/well in 96-well plates (Nunc-Immuno plate Maxisorp). Coated plates were covered and incubated at 4°C overnight. Plates were then washed 4 times with 0.05% Tween (Tween 20) in PBS, and blocked with 1% BSA/1% goat serum in PBS, 200 μl/well. After incubation with the blocking buffer at 37°C for 2 hours, the plates were washed 4 times with 0.05% Tween in PBS and tapped dry on paper towels.

B. Incubation with test samples and secondary antibody

Test samples were diluted to the proper concentrations in sample buffer, which contained 1% BSA/1% goat serum/0.05% Tween in PBS. A standard curve was prepared with a chimeric antibody (with a human Fc), the concentration of which was known. To prepare a standard curve, serial dilutions are made in the sample buffer to give a standard curve ranging from 125 ng/ml to 3.9 ng/ml. The diluted samples and standards were added to the plate, 100 μl/well, and the plate incubated at 37°C for 2 hours.

After incubation, the plate was washed 8 times with 0.05% Tween in PBS. To each well was then added 100 μl of the secondary antibody, the horse radish peroxidase (HRP) -conjugated anti-human IgG (Jackson Immuno Research), diluted around 1:120,000 in the sample buffer. The exact dilution of the secondary antibody had to be

determined for each lot of the HRP-conjugated anti-human IgG. After incubation at 37°C for 2 hours, the plate was washed 8 times with 0.05% Tween in PBS.

C. Development

The substrate solution was added to the plate at 100 μl/well. The substrate solution was prepared by dissolving 30 mg of o-phenylenediamine dihydrochloride (OPD) (1 tablet) into 15 ml of 0.025 M citric acid/0.05M Na_2HPO_4 buffer, pH 5, which contained 0.03% of freshly added H_2O_2. The color was allowed to develop for 30 minutes at room temperature in the dark. The developing time was subject to change, depending on lot to lot variability of the coated plates, the secondary antibody, etc. The color development in the standard curve was observed to determine when to stop the reaction. The reaction was stopped by adding 4N H_2SO_4, 100 μl/well. The plate was read by a plate reader, which was set at both 490 nm and 650 nm and programmed to subtract off the background OD at 650 nm from the OD at 490 nm.

The anti-hu kappa ELISA followed the same procedure as described above, except that the secondary antibody used was horse radish peroxidase-conjugated goat anti-hu kappa (Southern Biotechnology Assoc. Inc., Birmingham, AL), used at 1:4000 dilution.

The procedure for the anti-muFc ELISA was also similar, except that ELISA plates were coated with AffiniPure goat anti-murine IgG (H+L) (Jackson Immuno Research) at 5 μg/ml in PBS, and 100 μl/well; and the secondary antibody was horse radish peroxidase-conjugated goat anti-muIgG, Fcγ (Jackson ImmunoResearch), used at 1:5000 dilution.

1C. Cloning of the KS antigen (KSA, EpCAM) and expression of the soluble form as human EpCAM-murine Fc

Messenger RNA (MRNA) was prepared from LnCAP cells using Dynabeads mRNA Direct Kit (Dynal, Inc., Lake Success, NY) according to the manufacturer's instructions. After first strand cDNA synthesis with oligo(dT) and reverse transcriptase, full length cDNA encoding epithelial cell adhesion molecule (also known as KS antigen

or KSA), was cloned by polymerase chain reaction (PCR). The sequences of the PCR primers were based on the published sequence described in Perez and Walker (1989) J. Immunol. 142:3662-3667. The sequence of the sense primer is TCTAGAGCAGC**ATG**GCGCCCCCGCA (SEQ ID NO: 27), and the sequence of the nonsense primer is CTCGAG**TTA**TGCATTGAGTTCCCT (SEQ ID NO: 28), where the translation initiation codon and the anti-codon of the translation stop codon are denoted in bold, and the restriction sites XbaI (TCTAGA) and XhoI (CTCGAG) are underlined. The PCR product was cloned and the correct KSA sequence was confirmed by sequencing several independent clones. The cDNA sequence of the KSA from LnCAP was essentially identical to the published sequence of KSA from UCLA-P3 cells (Perez and Walker, 1989). However, at amino acid residue number 115, the nucleotide sequence from LnCAP was ATG rather than ACG (Met instead of Thr), and at amino acid residue number 277, the nucleotide sequence from LnCAP was ATA rather than ATG (Ile instead of Met).

Binding of KS-1/4 antibody to recombinant KSA was demonstrated by immunostaining. Surface expression of KSA was obtained by transfecting cells, e.g., CT26, B16, etc., with full length KSA in a suitable mammalian expression vector (pdCs, as described in U.S. Patent Number 5,541,087), followed by immunostaining with the KS-1/4 antibody. For the expression of KSA as a soluble antigen, the portion of the cDNA encoding the transmembrane domain of the KSA was deleted. To facilitate expression, detection, and purification, the soluble KSA was expressed as a KSA-muFc, the construction of which is described as follows. The 780 bp XbaI-EcoRI restriction fragment encoding the soluble KSA was ligated to the AflII-XhoI fragment encoding the muFc (U.S. Patent Number 5,726,044) via a linker-adaptor:

```
5'   AA TTC TCA ATG CAG GGC        3'  (SEQ ID NO:  29)

3'        G AGT TAC GTC CCG AAT T  5'  (SEQ ID NO:  30)
```

The XbaI-XhoI fragment encoding soluble KSA-muFc was ligated to the pdCs vector. The resultant expression vector, pdCs-KSA-muFc, was used to transfect cells and stable clones expressing KSA-muFc were identified by anti-muFc ELISA.

1D. Measurement of Antigen Binding

KSA-muFc in conditioned medium was first purified by Protein A chromatography according to supplier's protocol (Repligen, Cambridge, MA). Purified KSA-muFc was used to coat 96-well plates (Nunc-Immuno plate, Maxisorp) at 5 μg/ml in PBS, and 100 μl/well. The assay was similar to the ELISA procedure described in Example 1B. Briefly, coated plates were covered and incubated at 4°C overnight. Plates then were washed and blocked. Test samples were diluted to the proper concentrations in the sample buffer, added to the plate at 100 μl/well, and the plate was incubated at 37°C for 1 hour. After incubation, the plate was washed 8 times with 0.05% Tween in PBS. To each well was then added 100 μl of the secondary antibody, the horse radish peroxidase-conjugated anti-human IgG (Jackson Immuno Research), diluted around 1:120,000 in the sample buffer. The plate was then developed and read as described in Example 1B.

1E. Measurement of on-rates and off-rates of KS-1/4 antibodies from EpCAM using a Biacore assay.

The affinity of KS-1/4 and KS-IL2 molecules for the antigen EpCAM were measured by surface plasmon resonance analysis of the antibody-antigen interaction, using a Biacore machine (Biacore International AB, Uppsala, Sweden). EpCAM-murineFc was coupled to a CM5 sensor chip using an amine coupling protocol supplied by the manufacturer. KS-1/4 and KS-IL2 at concentrations varying between 25 nm and 200 nM were then passed over the chip, whereby binding to the chip was observed. Using the built-in curve-fitting routines of the Biacore software, the on-rate, off-rate, association and dissociation constants were calculated.

1F. Measurement of binding affinities of KS-1/4 antibodies using cell lines expressing EpCAM

Purified KS-1/4 antibodies were iodinated with ^{125}I using standard techniques, and increasing concentrations of labeled protein were incubated with the EpCAM-positive cell line PC-3. Saturation binding curves were generated and the dissociation constants were determined by Scatchard analysis.

Example 2. Cloning of cDNAs encoding V_H and V_K of mouse KS-1/4 and construction of vector for the expression of KS-1/4 hybridoma-derived antibody

Messenger RNA prepared from the mouse KS-1/4-expressing hybridoma (obtained from R. Reisfeld, Scripps Research Institute) was reverse transcribed with oligo(dT) and then used as templates for PCR to amplify the sequences encoding the variable region of the heavy chain (V_H) and the variable region of the light chain (V_K). The PCR primers were designed based on published sequences (Beavers *et al.*, ibid.). The PCR primers for V_H had the following sequences:

V_H forward primer (5') GACTCGAGCCCAAGTCTTAGACATC (3') (SEQ ID NO: 31)

V_H reverse primer (5') CAAGCTTACCTGAGGAGACGGTGACTGACGTTC (3'), (SEQ ID NO: 32)

where the CTCGAG and AAGCTT sequences represent the XhoI and HindIII restriction sites, respectively, used for ligating the V_H into the expression vector (see below); and the TAC in the reverse primer would introduce GTA, the splice donor consensus sequence, in the sense strand of the PCR product.

The PCR primers for V_K had the following sequences:

V_K forward primer (5') GATCTAGACAAG**ATG**GATTTTCAAGTG (3') (SEQ ID NO: 33)

V_K reverse primer (5') GAAGATCTTACGTTTTATTTCCAGCTTGG (3') (SEQ ID NO: 34)

where the TCTAGA and AGATCT sequences represent the XbaI and BglII restriction sites, respectively, used for ligating the V_K into the expression vector (see below); ATG is the translation initiation codon of the light chain; and the TAC in the reverse primer would introduce GTA, the splice donor consensus sequence, in the sense strand of the PCR product.

The PCR products encoding the V_H and V_K of the mouse KS-1/4 antibody were cloned into pCRII vector (Invitrogen, Carlsbad, CA). Several V_H and V_K clones were sequenced and the consensus sequence of each determined. The V_H and V_K sequences were inserted in a stepwise fashion into the expression vector pdHL7. The ligations took advantage of the unique XhoI and HindIII sites for the V_H, and the unique XbaI and BglII/BamHI sites for the V_K (the unique BglII in the V_K insert and the unique BamHI in the vector have compatible overhangs). The resultant construct is called pdHL7-hybridoma chKS-1/4, which already contained transcription regulatory elements and human Ig constant region sequences for the expression of chimeric antibodies (Gillies *et al.* (1989) J. Immunol. Methods 125:191).

The expression vector pdHL7 was derived from pdHL2 [Gillies *et al.* (1991) Hybridoma 10:347-356], with the following modifications: in the expression vector pdHL2, the transcriptional units for the light chain and the heavy chain-cytokine consisted of the enhancer of the heavy chain immunoglobulin gene and the metallothionein promoter. In pdHL7, these two transcriptional units consisted of the CMV enhancer-promoter [Boshart *et al.* (1985) Cell 41:521-530]. The DNA encoding the CMV enhancer-promoter was derived from the AflIII-HindIII fragment of the commercially available pcDNAI (Invitrogen Corp., San Diego, CA).

Example 3. Expression studies of murine KS-1/4 antibodies

This example discusses expression studies performed using an antibody expression plasmid encoding the V region sequences disclosed in U.S. Patent No. 4,975,369.

3A. Plasmid Construction

To directly compare the chimeric antibodies encoded by the Hybridoma KS-1/4 sequence and those sequences described in U.S. Patent No. 4,975,369, the cDNA encoding the VH sequence described in U.S. Patent No. 4,975,369 was synthesized. This was then ligated into the pdHL7 expression vector already containing the V_K of KS-1/4.

In order to construct the V_H sequence described in U.S. Patent No. 4,975,369, an NdeI-HindIII fragment encoding part of the V_H sequence was obtained by total chemical synthesis. Overlapping oligonucleotides were chemically synthesized and ligated. The ligated duplex was then subcloned into a XbaI-HindIII pBluescript vector (Stratagene, LaJolla, CA).

This DNA encodes the protein sequence IQQPQNMRTM of U.S. Patent No. 4,975,369. Immediately 3' to the coding sequence is the splice donor site beginning with gta. The ctag at the 5' end of the top strand is the overhang for the XbaI cloning site. The XbaI site was created only for cloning into the polylinker of the pBluescript vector. It was followed immediately by the NdeI restriction site (CATATG). The agct at the 5' end of the bottom strand is the overhang of the HindIII cloning site. This HindIII sticky end is later ligated to the HindIII site in the intron preceding the $C\gamma1$ gene [Gillies *et al.* (1991) Hybridoma 10:347-356].

After sequence verification, the NdeI-HindIII restriction fragment was isolated. This, together with the XhoI-NdeI fragment encoding the N-terminal half of V_H, was then ligated to the XhoI-HindIII digested pdHL7 expression vector containing the V_K of KS-1/4. The resultant construct, pdHL7-'369 chKS-1/4, contained the V_K and V_H described in U.S. Patent No. 4,975,369 (referred to as US4,975,369 chKS-1/4).

<u>3B. Comparison of hybridoma chKS-1/4 and US4,975,369 chKS-1/4 antibodies</u>

The plasmid DNAs pdHL7-hybridoma chKS-1/4 and pdHL7-'369 chKS-1/4 were introduced in parallel into human kidney 293 cells by the calcium phosphate coprecipitation procedure mentioned above. Five days post-transfection, the conditioned media were assayed by anti-huFc ELISA and kappa ELISA (see Example 1 for ELISA procedures) and the results are summarized in Table 1.

Table 1.

Antibody	huFc ELISA	Kappa ELISA
Hybridoma chKS-1/4	254 ng/mL	200 ng/mL
US4,975,369 chKS-1/4	14 ng/mL	0 ng/mL

The results indicated that hybridoma chKS-1/4 was expressed and secreted normally, and that the secreted antibody consisted of roughly equimolar amounts of heavy and light chains, within the accuracies of the two different ELISAs. On the other hand, only a low level of heavy chain was detected in the conditioned medium for the US4,975,369 chKS-1/4 antibody, and no kappa light chain was associated with it.

Western blot analysis was performed on the total cell lysates and the conditioned media of the two transiently transfected cell lines. The procedures for Western blot analysis were as described in (Sambrook *et al.* (1989), *supra*). In order to analyze the total cell lysates, the transfected cells were lysed, centrifuged to remove the debris, and the lysate from the equivalent of 5×10^5 cells applied per lane. To analyze the conditioned media, the protein product from 300 μL of the conditioned medium was first purified by Protein A Sepharose chromatography prior to SDS-PAGE under reducing conditions. After Western blot transfer, the blot was hybridized with a horseradish peroxidase-conjugated goat anti-human IgG, Fcγ (Jackson ImmunoResearch), used at 1:2000 dilution.

The Western blot transfer showed that under the conditions used, the heavy chain was detected in both the conditioned media and the lysed cells of the transfection with pdHL7-hybridoma chKS-1/4. This result indicates that the heavy chain of the chKS-1/4 antibody was produced in the cells and secreted efficiently (together with the light chain). On the other hand, the heavy chain from the transfection with pdHL7-'369 chKS-1/4 was detected only in the cell lysate but not in the conditioned media. This result indicated that although a comparable level of heavy chain was produced inside the cell, it was not secreted. This finding was consistent with the ELISA data, which showed that there was

no kappa light chain associated with the small amount of secreted heavy chain in the US4,975,369 chKS-1/4 antibody. It is understood that immunoglobulin heavy chains typically are not normally secreted in the absence of immunoglobulin light chains [Hendershot *et al.* (1987) Immunology Today 8:111].

In addition to the foregoing, NS/0 cells were transfected by electroporation with the plasmids pdHL7-Hybridoma chKS-1/4 and pdHL7-US4,975,369 chKS-1/4 in parallel. Stable clones were selected in the presence of 100 nM MTX, as described in Example 1, and the conditioned media of the MTX-resistant clones in 96-well plates was assayed by anti-huFc ELISA, as described in Example 1. The results are summarized in Table 2.

Table 2

Antibody	Total number of clones screened	Mode*	Highest level of expression*
Hybridoma chKS-1/4	80	0.1-0.5 μg/mL (41)	10-50 μg/mL (4)
US4,975,369 chKS-1/4	47	0-10 ng/mL (36)	0.1-0.4 μg/mL (4)

(*The numbers in parentheses denote the number of clones in the mode or the number expressing the highest levels of product, as determined by anti-Fc ELISA.)

When screened at the 96-well stage, the majority of the clones obtained with the pdHL7-hybridoma chKS-1/4 construct produced about 100 ng/mL to 500 ng/mL of antibody, with the best clones producing about 10-50 μg/mL. On the other hand, the majority of the clones obtained with the pdHL7-'369 chKS-1/4 construct produced about 0 ng/mL to 10 ng/mL of antibody, with the best producing about 300-400 ng/mL. To examine the composition and binding properties of the US4,975,369 chKS-1/4 antibody, it was necessary to grow up the clones that produced at 300-400 ng/mL. Two of these clones were chosen for expansion. However, their expression levels were found to be very unstable. By the time the cultures were grown up to 200 mL, the expression levels of both clones had dropped to about 20 ng/mL, as assayed by anti-Fc ELISA. When the same conditioned media were assayed by the anti-kappa ELISA, no kappa light chain was detected, as was the case in transient expression in 293 cells.

The following experiment indicated that no detectable kappa light chain was associated with the US4,975,369 chKS-1/4 heavy chain. Briefly, 50 mL each of the conditioned media from each of the clones was concentrated by Protein A chromatography. The eluate were assayed by anti-Fc ELISA and anti-kappa ELISA. As a control, conditioned medium from a hybridoma chKS-1/4-producing clone was treated the same way and assayed at the same time. The ELISA results are summarized in Table 3.

Table 3

Antibody	huFc ELISA	Kappa ELISA
Hybridoma chKS-1/4	42 μg/mL	44 μg/mL
US4,975,369 chKS-1/4-clone 1	253 ng/mL	0 ng/mL
US4,975,369 chKS-1/4-clone 2	313 ng/mL	0 ng/mL

The results showed that there was indeed no detectable kappa light chain associated with the US4,975,369 chKS-1/4 heavy chain. Furthermore, the hybridoma chKS-1/4 antibody was shown to bind KS antigen at 10-20 ng/mL, whereas the US4,975,369 antibody from both clones and concentrated to 253 and 313 ng/mL, still did not bind KS antigen (see Example 9 for measurement of binding to KS antigen.)

Example 4. Expression and characterization of variant KS antibodies

Mutations that significantly lower the expression or the affinity of an antibody for a target molecule are expected to be less effective for therapeutic purposes in humans. Some approaches to reducing immunogenicity, such as "veneering," "humanization," and "deimmunization" involve the introduction of many amino acid substitutions, and may disrupt binding of an antibody to an antigen (see, e.g., U.S. Patent Nos. 5,639,641; and 5,585,089; and PCT Publication Nos. WO 98/52976; WO 00/34317). There is a need in the art for classes of antibody sequences that will bind to epithelial cell adhesion molecule, but which are distinct from the original mouse monoclonal antibodies that recognize this antigen.

Various combinations of KS-1/4 heavy and light chain variable ("V") regions were tested for their ability to be expressed, and for their ability to bind to EpCAM. These results are summarized in Tables 4-6 and described below.

Table 4. Sequences of KS-1/4 antibody heavy and light chain V regions.

Light chains:

```
                10        20        30        40        50        60
                 |         |         |         |         |         |
VK0     QILLTQSPAIMSASPGEKVTMTCSASSSVSYMLWYQQKPGSSPKPWIFDTSNLASGFPAR
VK1     QIVLTQSPASLAVSPGQRATITCSASSSVSYILWYQQKPGQPPKPWIFDTSNLASGFPSR
VK6     EIVLTQSPATLSLSPGERVTLTCSASSSVSYMLWYQQKPGQAPKLLIFDTSNLASGIPAR
VK7     QILLTQSPAIMSASPGERVTMTCSASSSVSYMLWYQQKPGSSPKPWIFDTSNLASGFPAR
VK8     EIVLTQSPATLSLSPGERVTLTCSASSSVSYMLWYQQKPGSSPKPWIFDTSNLASGFPAR

                70        80        90        100
                 |         |         |         |
VK0     FSGSGSGTSYSLIISSMEAEDAATYYCHQRSGYPYTFGGGTKLEIK (SEQ ID NO:  1)
VK1     FSGSGSGTSYTLTINSLEAEDAATYYCHQRSGYPYTFGGGTKVEIK (SEQ ID NO: 11)
VK6     FSGSGSGTDYTLTISSLEPEDFAVYYCHQRSGYPYTFGGGTKLEIK (SEQ ID NO:  7)
VK7     FSGSGSGTSYSLIISSMEPEDAATYYCHQRSGYPYTFGGGTKLEIK (SEQ ID NO:  8)
VK8     FSGSGSGTSYSLIISSMEAEDAATYYCHQRSGYPYTFGGGTKLEIK (SEQ ID NO:  9)
```

Heavy chains:

```
                10        20        30        40        50        60
                 |         |         |         |         |         |
VH0     QIQLVQSGPELKKPGETVKISCKASGYTFTNYGMNWVKQTPGKGLKWMGWINTYTGEPTY
VH1     QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVRQAPGKGLKWMGWINTYTGEPTY
VH2     QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVRQAPGKGLKWMGWINTYTGEPTY
VH2.5   QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVRQAPGKGLKWMGWINTYTGEPTY
VH6     QVQLVQSGAEVKKPGESVKISCKASGYTFTNYGMNWVRQAPGKGLEWMGWINTYTGEPTY
VH7     QIQLVQSGAEVKKPGETVKISCKASGYTFTNYGMNWVKQTPGKGLKWMGWINTYTGEPTY
VH369   QIQLVQSGPELKKPGETVKISCKASGYTFTNYGMNWVKQTPGKGLKWMGWINTYTGEPTY

                70        80        90        100       110
                 |         |         |         |         |
VH0     ADDFKGRFAFSLETSASTAFLQINNLRNE.DMATYFCVRFISKGDYWGQGTSVTVSS (SEQ ID NO:  2)
VH1     ADDFKGRFTITAETSTSTLYLQLNNLRSE.DTATYFCVRFMSKGDYWGQGTTVTVSS (SEQ ID NO: 21)
VH2     ADDFKGRFTITAETSTSTLYLQLNNLRSE.DTATYFCVRFISKGDYWGQGTTVTVSS (SEQ ID NO: 22)
VH2.5   ADDFKGRFTITAETSTSTLYLQLNNLRSE.DTATYFCVRFISKGDYWGTGTTVTVSS (SEQ ID NO: 19)
VH6     AQKFQGRVTISLDTSTSTAYLQLSSLRAE.DTAVYFCVRFISKGDYWGQGTSVTVSS (SEQ ID NO: 17)
VH7     ADDFKGRFAFSLETSTSTAFLQINNLRSE.DTATYFCVRFISKGDYWGQGTSVTVSS (SEQ ID NO: 18)
VH369   ADDFKGRFAFSLETSASTAFLQIqqpqnmrtMATYFCVRFISKGDYWGQGTSVTVSS (SEQ ID NO: 35)
```

Table 5. Sequences of KS-1/4 antibody variants and CDR3 heavy chain variants with single amino acid insertions.

```
VH2 partial seq.:      . . . ATYFCVRF I S K GDYWGQG. . . (amino acid residues
92-109 of SEQ ID NO:  22)
VH2.1:                 . . . ATYFCVRF IIS K GDYWGQG. . . (SEQ ID NO:  36)
VH2.2:                 . . . ATYFCVRF IVS K GDYWGQG. . . (SEQ ID NO:  37)
VH2.3:                 . . . ATYFCVRF I SAK GDYWGQG. . . (SEQ ID NO:  38)
VH2.4:                 . . . ATYFCVRF I S KTGDYWGQG. . . (SEQ ID NO:  39)
```

Table 6. Expression levels and binding activity of variant KS-1/4 antibodies.

Construct	Expression		EpCAM affinity	
	Transient (*) (in ng/mL)	Stable (*) (in μg/mL)	Relative binding (**)	Kd (nM)
Group 1				
VK0/VH0 (Hybridoma chKS-1/4)		10 – 50	1x	1.0×10^{-9}
VK0/VH'369 ('369 chKS-1/4)		0.1 – 0.4(***)	>>30x	
VK8/VH7 (Construct 3)		10 – 50		1.0×10^{-9}
VK6/VH6 (Construct 1)	300		n.d.	
VK7/VH7 (Construct 2)	30			
VK8/VH7-IL2		10 – 50		1.0×10^{-9}
VK1/VH1-IL2		10 – 50		7.9×10^{-9}
VK1/VH2-IL2		10 – 50		3.1×10^{-9}
Group 2				
VK8/VH7 (Construct 3; control)	1500		1x	
VK0/VH1	1500		8x	
VK1/VH7	1500		1x	
VK1/VH1	1500		2x	
VK1/VH2	1500		1x –2x	
VK1/VH1-IL2	1500		5x	
VK1/VH2-IL2	1500		1.5x	
VK1/VH2.5-IL2	1500		3x – 4x	
Group 3				
VK8/VH7-IL2 (control)	760		1x	
VK1/VH1-IL2	350		2x	
VK1/VH2.1-IL2	290		>10x	
VK1/VH2.2-IL2	270		>10x	
VK1/VH2.3-IL2	190		7x	
VK1/VH2.4-IL2	210		3x	

(*) Routinely achievable levels.
(**) "Relative Binding" is expressed as the fold-increase in protein concentration required to reach an equivalent level of binding. Thus, a larger number reflects a lower affinity for EpCAM.
(***) Kappa light chain was not detectable by ELISA (equivalent to background); therefore, functional antibodies were not expressed.
(****) n.d. = not detectable
In Group 2 and Group 3, the relative binding activity of each protein was normalized to the control shown in the first line for that group. The ELISA assay is primarily a reflection of off-rates, based on amount of protein bound after several rounds of washes. It is used as a rapid screen to rule out poor binders, but is not a

precise measure of affinity. In Group 3, VH2 variants VH2.1 – VH2.4 were compared with VH1 to determine if amino acid insertions might result in improved relative binding.

The sequences are related as follows. As described in the examples, the VH0 and VK0 sequences were derived from PCR amplification from a hybridoma cell line that expresses the original mouse-derived KS-1/4 (SEQ ID NO: 1 and SEQ ID NO: 2). VH-'369 is the VH sequence disclosed in U.S. Patent No. 4,975,369. Sequences VH1, VH2, VH2.1-2.4 VK1, and VK2 were derived either using deimmunization technology where potential T cell epitopes are eliminated or weakened by introduction of mutations that reduce binding of a peptide epitope to an MHC Class II molecule, or by changing non-human T cell epitopes so that they correspond to human self-epitopes that are present in human antibodies. The design of these constructs is further described and analyzed below. Constructs of Table 6 were generated by transfecting mammalian cells with combinations of nucleic acids that expressed the corresponding heavy and light chain V regions. Sequences VH6, VH7, VK6, VK7, and VK8 were generated by changing surface residues of the hybridoma KS-1/4 to human counterparts as described below, with the purpose of removing potential human B cell epitopes. Constructs 1 through 3 were generated by transfecting mammalian cells with combinations of nucleic acids that expressed heavy and light chain V regions VH6, VH7, VK6, VK7, and VK8 as described in Table 4 and below.

4A. Characterization of KS antibodies with fewer human T cell epitopes

Sequences VH2.1-VH2.5 were made to test whether certain amino acid insertions and substitutions in the region of the KS-1/4 heavy chain CDR3 could be tolerated. Expression vectors for the light and heavy chain combinations VK0/VH1, VK1/VH7, VK1/VH1, VK1/VH2, VK1/VH1-IL2, VK1/VH2-IL2, and VK1/VH2.5-IL2 were constructed and the corresponding antibodies and antibody-IL2 fusion proteins expressed and tested according to methods described in the preceding examples.

Specifically, sequences VH1, VH2, VK1, and VK2 were obtained by total chemical synthesis. For each of these sequences, a series of overlapping oligonucleotides that span the entire coding and complementary strands of these regions were chemically

synthesized, phosphorylated, and ligated. The ligated duplex molecules were then amplified by PCR with appropriate primers to the fragment ends, introduced into pCRII vector (Invitrogen, Carlsbad, CA) and the sequences verified. These DNA fragments were then introduced into the expression vector pdHL7 at appropriate sites to generate the complete heavy ("H") chain and light ("L") chain, respectively.

Sequence VH2.5 was derived from VH2 by the modification of a single codon to obtain a Thr rather than a Gln at position 108 (Table 4), using standard molecular biology techniques.

The antibodies were tested by ELISA (Table 6) and using surface plasmon resonance (Biacore machine and software) to compare their ability to bind to EpCAM. Results of the ELISA experiments were considered to reflect primarily off-rate and not on-rate, and to be generally less precise, such that a poor ELISA result was generally used to exclude certain constructs from further consideration. However, antibodies that showed good binding by the ELISA test needed to be characterized further.

Results of the surface plasmon resonance analysis were as follows:

Fusion Protein	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_D (M)
VK8/VH7-IL2	3.1×10^5	3.2×10^{-4}	1.0×10^{-9}
VK1/VH2-IL2	1.7×10^5	5.3×10^{-4}	3.1×10^{-9}
VK1/VH1-IL2	2.8×10^5	2.2×10^{-3}	7.9×10^{-9}

Because the off-rate of VK1/VH1-IL2 was much faster than for VK1/V2-IL2 or VK8/VH7-IL2, VK1/VH1-IL2 was considered to be a less useful fusion protein.

Considering that VK1/VH1-IL2 and VK1/VH1-IL2 differ only by the methionine/isoleucine difference at V_H position 100 in CDR3, the enhanced off-rate of VK1/VH1-IL2 compared to VK1/VH2-IL2 suggests that this position makes a

hydrophobic contact with EpCAM, and that the slightly longer methionine side-chain makes a less effective contact. In the field of protein-protein interactions, it is generally thought that hydrophobic interactions play a major role in determining off-rates but a much less significant role in determining on-rates.

4B. Characterization of KS-1/4 variants with single amino acid insertions

The importance of the CDR3 sequence in the heavy chain V region for the affinity of the KS antibody to EpCAM was determined with a series of variants that contained an amino acid insertion or substitution in this region. Sequences VH2.1, VH2.2, VH2.3, and VH2.4 were generated by manipulation of an expression vector encoding VH2 and VK1 using standard recombinant DNA techniques. The resulting expression vectors were transfected into NS/0 cells and secreted antibody proteins purified as described in preceding examples.

It was found that the VH1 variant was suboptimal compared to the VH2 variant, indicating that the isoleucine in CDR3 could not be substituted with methionine. The next goal was to test whether insertion of an amino acid in CDR3 could yield a KS-1/4 heavy chain V region with better binding characteristics than VH1. The data in Table 6 compare the binding of VK1/VH2.1, VK1/VH2.2, VK1/VH2.3, and VK1/VH2.4, with VK1/VH1. It was found that none of the constructs with an amino acid insertion in the KS-1/4 V_H CDR3 showed improved antigen binding compared to VH1, rather, antigen binding activity of the insertion mutants was either somewhat decreased or profoundly decreased.

These results indicate that insertion of amino acids in CDR3 generally is deleterious to the antigen binding activity of KS-1/4 heavy chain V regions. When this data is analyzed, some general conclusions emerge. Specifically, the segment of KS-1/4 V_H amino acid at positions 84 to 108, consisting of the amino acids Asn-Asn-Leu-Arg-Asn-Glu-Asp-Met-Ala-Thr-Tyr-Phe-Cys-Val-Arg-Phe-Ile-Ser-Lys-Gly-Asp-Tyr-Trp-Gly-Gln, is important for KS-1/4 antigen binding. This segment includes a framework segment, Asn-Asn-Leu-Arg-Asn-Glu-Asp-Met-Ala-Thr-Tyr-Phe-Cys-Val-Arg, which is generally tolerant to single and multiple amino acid substitutions, but not tolerant to

amino acid insertions, which may have a deleterious effect on expression and assembly. In addition, the data suggests that for the amino acids at positions 86, 91, 93, 94, and 95, it is preferable to have hydrophobic amino acids for an antibody that is efficiently expressed and binds to EpCAM.

Insertion of an amino acid in the V_H CDR3 segment, consisting of Phe-Ile-Ser-Lys-Gly-Asp-Tyr, is generally deleterious to the EpCAM antigen-binding function of a KS-1/4 antibody, although some insertions can be tolerated with only partial loss of activity. Similarly, substitution of these positions is also generally deleterious to binding of the EpCAM antigen, although some insertions can be tolerated with only partial loss of activity.

4C. Construction of active derivatives of KS-1/4 antibodies with mouse surface residues converted to their human counterparts

Antibodies were prepared by substituting amino acids within the KS-1/4 antibody with amino acids commonly found in human antibodies in order to minimize the immunogenicity of the mouse-derived V regions. Preferred KS derivatives also retained specific binding affinity for human EpCAM.

Construct 1. It was found that the KS-1/4 light chain most closely resembled human consensus subgroup III, and the heavy chain most closely resembled subgroup I. Based on these similarities, a conceptual sequence consisting of the human consensus subgroup amino acids and KS-1/4-derived CDRs and non-consensus amino acids was generated. For this and the following constructs a three-dimensional model was generated using a Silicon Graphics Workstation and BioSym molecular modeling software.

Inspection of the three-dimensional model revealed that certain human-derived amino acids were close to the CDRs and were likely to influence their conformation. Based on this analysis, in the light chain, human Ser22, Arg44, and Phe66 were changed back to Thr, Lys, and Tyr, respectively. In the heavy chain, it was believed such changes were unnecessary.. In the final design for Construct 1, the light chain had 18 human

amino acids not found in the mouse light chain, and the heavy chain had 22 human amino acids not found in the mouse heavy chain.

DNAs for expression of Construct 1 were created using synthetic oligonucleotides. The Construct 1 protein was efficiently expressed but was found to be more than 10-fold less active in an EpCAM binding assay.

Construct 2. A less aggressive approach was then taken, by which only the following changes were introduced:

Light chain: K18R, A79P

Heavy chain: P9A, L11V, A76T, N88S, M91T

DNAs for expression of Construct 2 were created using synthetic oligonucleotides and standard recombinant DNA techniques. The Construct 2 protein was not efficiently expressed. It was further found that the combination of Construct 2 light chain and mouse KS-1/4 heavy chain was not efficiently expressed, while the combination of Construct 2 heavy chain and mouse KS-1/4 light chain was efficiently expressed. Thus, the expression defect appeared to lie in the Construct 2 light chain.

Construct 3. Based on the apparent expression defect in the Construct 2 light chain, a new light chain was constructed by fusing the N-terminal portion of the light chain of Construct 1 with the C-terminal portion of the mouse light chain. The KpnI site, which encodes the amino acids at positions 35 and 36, was used. When this light chain was combined with the Construct 2 heavy chain, efficient expression and no significant loss of binding was observed.

Because Construct 3 resulted in an antibody with superior properties in terms of protein expression and affinity for the antigen when compared to Construct 1 or 2, DNA sequences of Construct 3 were inserted into pdHL7s-IL2, resulting in pdHL7s-VK8/VH7-IL2, which is disclosed as SEQ ID NO: 40. For expression purposes, this plasmid DNA was electroporated into mouse myeloma cells NS/0 to produce a stably transfected cell line as described in Example 1A. Culture medium taken from stable

clones was then assayed for antibody expression in an ELISA coated with human Fc, as described in Example 1B. The amino acid sequences of the heavy and light chain for this antibody fusion protein are shown in SEQ ID NO: 41 and SEQ ID NO: 42, respectively.

In addition, the binding of iodinated VK8/VH7 and VK8/VH7-IL2 to EpCAM expressed on the surface of PC-3 tumor cells was compared to binding of iodinated VK0/VH0-IL2, using methods described in Example 1F. Within experimental error, essentially identical binding affinities were found for VK8/VH7 and VK0/VH0, and for VK8/VH7-IL2 and VK0/VH0-IL2.

4D. Structure-function relationships useful in constructing active KS-1/4 antibodies

Taken together, the antigen binding activities of KS-1/4 antibodies and fusion proteins with the disclosed V region sequences provide guidance in designing sequences of KS-1/4 antibodies to EpCAM, as well as for proper expression and secretion of KS-1/4 antibodies. In particular, the KS-1/4 heavy and light chain V regions can tolerate multiple amino acid substitutions and retain activity, provided that these amino acid substitutions are outside the CDRs. The KS-1/4 heavy and light chain V regions do not generally appear to tolerate amino acid insertions, especially within CDRs or in framework regions between CDRs.

For example, if the hybridoma KS-1/4 sequence is taken to be a starting, "wild-type" sequence, the data indicate that the heavy chain V region can tolerate amino acid substitutions at positions 9, 11, 16, 17, 38, 40, 69, 70, 71, 72, 76, 79, 80, 83, 88, 91, and 111 with little or no loss of activity. Similarly, the light chain can tolerate amino acid substitutions at positions 1, 3, 10, 11, 12, 13, 17, 18, 19, 21, 41, 42, 59, 71, 73, 75, 77, and 103 with little or no loss of activity. These changes are outside the CDRs of KS-1/4 heavy and light chain V regions. The 17 clearly acceptable heavy chain amino acid substitutions represent about 21% of the amino acid positions outside the CDRs, and about 68% of the amino acid positions outside the CDRs for which an amino acid substitution was attempted. Similarly, the eighteen clearly acceptable light chain amino acid substitutions represent about 23% of the amino acid positions outside the CDRs, and about 72% of the amino acid positions outside the CDRs for which an amino acid

substitution was attempted. There were only two examples of an amino acid substitution outside of a CDR that resulted in a significantly less useful protein: the substitution Ala79Pro in the light chain, which appeared to have a negative impact on expression; and the substitution Q108T in the heavy chain, which had a negative impact on antigen binding. Thus, an amino acid substitution can be introduced into a KS-1/4 antibody heavy chain or light chain sequence outside of a CDR, and there is a high probability that the substitution will result in an active protein.

Mutations involving the substitution of an amino acid in a CDR often have a negative impact on antigen binding. For example, the substitution I100M in the heavy chain reduces binding by about 8-fold. Mutations that involve the insertion of an amino acid generally have a negative impact on the utility of a KS-1/4 sequence. For example, the VH-'369 heavy chain V region is unable to assemble into a proper antibody with a light chain, as described herein. The VH2.1 to 2.4 mutations have an insertion of an amino acid in CDR3 of the heavy chain V region, and each of these mutations has a negative impact on antigen binding.

Example 5. Immunogenicity of a KS Antibody (Construct 3)-IL2 Fusion Protein in Humans

In a human clinical trial, twenty two patients received one or more treatment regimes, with each treatment regime comprising three consecutive daily 4-hour intraveous infusions of KS antibody (Construct 3)-IL2. Each treatment regime was separated by a month (Weber *et al.* (2001). Proc. Am. Soc. Clin. Oncology 20:259a.). Serum samples were harvested from each patient before and after each treatment regime and tested for antibody reactivity against the whole KS Antibody (Construct 3)-IL2 molecule or the Fc-IL2 component (without the Fv region). No reactivity was observed in any of the pre-immune sera. The results indicated that only 4 patients experienced any significant immune response against either the Fv regions alone, or both the Fv regions and the Fc-IL2 component. Furthermore, these responses did not appear to be boosted upon subsequent exposure to huKS-IL2.

It is believed that the use of the antibody-IL2 fusion protein constitutes a particularly stringent test of the immunogenicity of the V region, because the interleukin-2 moiety has an adjuvant effect. Accordingly, the results indicate that the KS Antibody (Construct 3) may be administered to humans with only a small number of recipients apparently developing an antibody response to the KS antibody (Construct 3)-IL2 fusion protein. These results are particularly encouraging in view of the fact that the KS antibody (Construct 3) contains a variable region that is almost entirely murine in origin but with a few amino acid residues replaced with the corresponding human amino acid residues.

EQUIVALENTS

The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative rather than limiting on the invention described herein. The scope of the invention is thus indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are intended to be embraced therein.

INCORPORATION BY REFERENCE

The disclosure of each of the patent documents and scientific publications disclosed herein, are incorporated by reference into this application in their entirety.

What is claimed is:

1. A recombinant anti-EpCAM antibody, wherein the antibody comprises an amino acid sequence defining an immunoglobulin light chain framework region selected from the group consisting of:

(i) amino acid residues 1-23 of SEQ ID NO: 5, wherein Xaa1 is Q or E, Xaa3 is L or V, Xaa10 is I or T, Xaa11 is M or L, Xaa13 is A or L, Xaa18 is K or R, or Xaa21 is M or L, provided that at least one of the amino acid residues at positions Xaa1, Xaa3, Xaa10, Xaa11, Xaa13, Xaa18, or Xaa21 is not the same as the amino acid at the corresponding position in SEQ ID NO: 1;

(ii) amino acid residues 34-48 of SEQ ID NO: 5, wherein Xaa41 is S or Q, Xaa42 is S or A, Xaa45 is P or L, or Xaa46 is W or L, provided that at least one of the amino acid residues at positions Xaa41, Xaa42, Xaa45, or Xaa46 is not the same as the amino acid at the corresponding position in SEQ ID NO: 1; and

(iii) amino acid residues 56-87 of SEQ ID NO: 5, wherein Xaa57 is F or I, Xaa69 is S or D, Xaa71 is S or T, Xaa73 is I or T, Xaa77 is M or L, Xaa79 is A or P, Xaa82 is A or F, or Xaa84 is T or V, provided that at least one of the amino acid residues at positions Xaa57, Xaa69, Xaa71, Xaa73, Xaa77, Xaa79, Xaa82, or Xaa84 is not the same as the amino acid at the corresponding position in SEQ ID NO: 1.

2. A recombinant anti-EpCAM antibody, wherein the antibody comprises an amino acid sequence defining an immunoglobulin heavy chain framework region selected from the group consisting of:

(i) amino acid residues 1-25 of SEQ ID NO: 6, wherein Xaa2 is I or V, Xaa9 is P or A, Xaa11 is L or V, or Xaa17 is T or S, provided that at least one of the amino acid residues at positions Xaa2, Xaa9, Xaa11 or Xaa17 is not the same as the amino acid at the corresponding position in SEQ ID NO: 2;

(ii) amino acid residues 36-49 of SEQ ID NO: 6, wherein Xaa38 is K or R, Xaa40 is T or A, or Xaa46 is K or E, provided that at least one of the amino acid residues at

positions Xaa38, Xaa40, Xaa46 is not the same as the amino acid at the corresponding position in SEQ ID NO: 2;

(iii) amino acid residues 67-98 of SEQ ID NO: 6, wherein Xaa68 is F or V, Xaa69 is A or T, Xaa70 is F or I, Xaa73 is E or D, Xaa76 is A or T, Xaa80 is F or Y, Xaa83 is I or L, Xaa84 is N or S, Xaa85 is N or S, Xaa88 is N, A or S, Xaa91 is M or T, or Xaa93 is T or V, provided that at least one of the amino acid residues at positions Xaa68, Xaa69, Xaa70, Xaa73, Xaa76, Xaa80, Xaa83, Xaa84, Xaa85, Xaa88, Xaa91 or Xaa93 is not the same as the amino acid at the corresponding position in SEQ ID NO: 2; and

(iv) amino acid residues 106-116 of SEQ ID NO: 6, wherein Xaa108 is Q or T.

3. The recombinant antibody of claim 1, wherein said light chain framework region is selected from the group consisting of:

(i) amino acid residues 1-23 of SEQ ID NO: 8; and

(ii) amino acid residues 1-23 of SEQ ID NO: 9.

4. The recombinant antibody of claim 3, wherein said light chain comprises amino acids 1-106 of SEQ ID NO: 9.

5. The recombinant antibody of claim 2, wherein said heavy chain framework region is selected from the group consisting of:

(i) amino acid residues 1-25 of SEQ ID NO: 18; and

(ii) amino acid residues 67-98 of SEQ ID NO: 18.

6. The recombinant antibody of claim 5, wherein said heavy chain comprises amino acids 1-116 of SEQ ID NO: 18.

7. A recombinant anti-EpCAM antibody comprising light chain amino acid residues 1-106 of SEQ ID NO: 9 and heavy chain amino acid residues of SEQ ID NO: 18.

8. The recombinant antibody of claim 1 or 2 wherein said antibody has a Kd for EpCAM of at least 10^{-8} M.

9. The recombinant antibody of claim 1 or 2 wherein said antibody comprises a cytokine.

10. The recombinant antibody of claim 9 wherein said cytokine is IL-2.

11. The recombinant antibody of claim 1 wherein said antibody comprises an amino acid sequence selected from the group consisting of:

(i) amino acid residues 24-31 of SEQ ID NO: 1;

(ii) amino acid residues 49-55 of SEQ ID NO: 1; and

(iii) amino acid residues 88-96 of SEQ ID NO: 1.

12. The recombinant antibody of claim 2 wherein said antibody comprises an amino acid sequence selected from the group consisting of:

(i) amino acid residues 26-35 of SEQ ID NO: 2;

(ii) amino acid residues 50-62 of SEQ ID NO: 2; and

(iii) amino acid residues 101-105 of SEQ ID NO: 2.

13. An expression vector encoding an antibody of claim 1 or 2.

14. An expression vector encoding the antibody of claim 7.

15. An expression vector having a nucleotide sequence set forth in SEQ ID NO: 32.

16. A method of treating a human patient having a disease associated with EpCAM over-expression, said method comprising the step of administering an antibody of claim 1 or 2 to a patient.

17. The method of claim 16, wherein said antibody further comprises a cytokine.

18. The method of claim 17, wherein said antibody is administered as an antibody-cytokine fusion protein.

IA

	FR1 (10, 20) / CDR1 (30)
‹S-1/4 VK0 (Light) (SEQ ID NO: 1)	Q I L L T Q S P A I M S A S P G E K V T M T C S A S S S V S Y
ct 1 VK6 (Light) (SEQ ID NO: 7)	E I V L T Q S P A T L S L S P G E R V T L T C S A S S S V S Y
ct 2 VK7 (Light) (SEQ ID NO: 8)	Q I L L T Q S P A I M S A S P G E R V T M T C S A S S S V S Y
ct 3 VK8 (Light) (SEQ ID NO: 9)	E I V L T Q S P A T L S L S P G E R V T L T C S A S S S V S Y
sus Sequence 3 VK (Light) (SEQ ID NO: 5)	X I X L T Q S P A X X S X S P G E X V T X T C S A S S S V S Y
'eneered (Light) (SEQ ID NO: 10)	Q I L L T Q S P A S L A V S P G Q R A T I T C S A S S S V S Y
mmunized VK1 (Light) (SEQ ID NO: 11)	Q I V L T Q S P A S L A V S P G Q R A T I T C S A S S S V S Y
mmunized VK2 (Light) (SEQ ID NO: 12)	Q I V L T Q S P A S L A V S P G Q R A T I T C S A S S S V S Y
mmunized VK3 (Light) (SEQ ID NO: 13)	Q I L L T Q S P A S L A V S P G Q R A T I T C S A S S S V S Y
mmunized VK4 (Light) (SEQ ID NO: 14)	Q I L L T Q S P A S L A V S P G Q R A T I T C S A S S S V S Y
mmunized VK5 (Light) (SEQ ID NO: 15)	Q I L L T Q S P A S L A V S P G Q R A T I T C S A S S S V S Y
mouse (Mo PT) (Light) (SEQ ID NO: 16)	Q I V L T Q S P A T L S A S P G E R V T I T C S A S S S V S Y
sus Sequence 1 VK (Light) (SEQ ID NO: 3)	X I X L T Q S P A X X X X S P G X X X T X T C S A S S S V S T

	FR1 (10, 20) / CDR1 (30)
‹S-1/4 VH0 (Heavy) (SEQ ID NO: 2)	Q I Q L V Q S G P E L K K P G E T V K I S C K A S G Y T F T N
ct 1 VH6 (Heavy) (SEQ ID NO: 17)	Q V Q L V Q S G A E V K K P G E S V K I S C K A S G Y T F T N
ct 3 VH7 (Heavy) (SEQ ID NO: 18)	Q I Q L V Q S G A E V K K P G E T V K I S C K A S G Y T F T N
sus Sequence 4 VH (Heavy) (SEQ ID NO: 6)	Q X Q L V Q S G X E X K K P G E X V K I S C K A S G Y T F T N
nmunized VH2.5 (Heavy) (SEQ ID NO: 19)	Q I Q L V Q S G P E L K K P G S S V K I S C K A S G Y T F T N
'eneered (Heavy) (SEQ ID NO: 20)	Q I Q L V Q S G P E L K K P G S S V K I S C K A S G Y T F T N
mmunized VH1 (Heavy) (SEQ ID NO: 21)	Q I Q L V Q S G P E L K K P G S S V K I S C K A S G Y T F T N
nmunized VH2 (Heavy) (SEQ ID NO: 22)	Q I Q L V Q S G P E L K K P G S S V K I S C K A S G Y T F T N
mmunized VH3 (Heavy) (SEQ ID NO: 23)	Q I Q L V Q S G P E L K K P G S S V K I S C K A S G Y T F T N
mmunized VH4 (Heavy) (SEQ ID NO: 24)	Q I Q L V Q S G P E L K K P G S S V K I S C K A S G Y T F T N
mmunized VH5 (Heavy) (SEQ ID NO: 25)	Q I Q L V Q S G P E L K K P G S S V K I S C K A S G Y T F T N
mouse (Mo PT) (Heavy) (SEQ ID NO: 26)	Q I Q L V Q S G P E L K K P G E T V K I S C K A S G Y T F T N
sus Sequence 2 VH (Heavy) (SEQ ID NO: 4)	Q X Q L V Q S G X E X K K P G X X V K I S C K A S G Y T F T N

Figure 1B

```
                FR2                       CDR2                               FR3
                  40                      50                      60                      70                 80
M L W Y Q Q K P G S S P K P W I F D T S N L A S G F P A R F S G S G S G T S Y S L I I S S M E A E
M L W Y Q Q K P G Q A P K L L I F D T S N L A S G I P A R F S G S G S G T D Y T L T I S S L E P E
M L W Y Q Q K P G S S P K P W I F D T S N L A S G F P A R F S G S G S G T S Y S L I I S S M E P E
M L W Y Q Q K P G S S P K P W I F D T S N L A S G F P A R F S G S G S G T S Y S L I I S S M E A E
M L W Y Q Q K P G X X P K X X I F D T S N L A S G X P A R F S G S G S G T X Y X L X I S S X E X E
M L W Y Q Q K P G Q P P K P W I F D T S N L A S G F P A R F S G S G S G T S Y T L T I N S L E A E
I L W Y Q Q K P G Q P P K P W I F D T S N L A S G F P S R F S G S G S G T S Y T L T I N S L E A E
M L W Y Q Q K P G Q P P K P W I F D T S N L A S G F P S R F S G S G S G T S Y T L T I N S L E A E
M L W Y Q Q K P G Q P P K P W I F D T S N L A S G F P S R F S G S G S G T S Y T L T I N S L E A E
M L W Y Q Q K P G Q P P K P W I F D T S N L A S G F P A R F S G S G S G T S Y T L T I N S L E A E
M L W Y Q Q K P G S S P K P W I Y D T S N L A S G F P A R F S G S G S G T S Y T L T I N S L E A E
M L W Y L Q K P G S S P K P W I F D T S N L A S G F P S R F S G S G S G T T Y S L I I S S L E A E
X L W Y X Q K P G X X P K X X I X D T S N L A S G X P X R F S G S G S G T X Y X L X I X S X E X E
```

```
                    FR2                          CDR2                                      FR3
                  40                      50                      60                      70                 8
Y G M N W V K Q T P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F A F S L E T S A S T A F
Y G M N W V R Q A P G K G L E W M G W I N T Y T G E P T Y A Q K F Q G R V T I S L D T S T S T A Y
Y G M N W V K Q T P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F A F S L E T S T S T A F
Y G M N W V X Q X P G K G L X W M G W I N T Y T G E P T Y A D X F X G R X X X S L X T S X S T A X
Y G M N W V R Q A P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F T I T A E T S T S T L Y
Y G M N W V K Q A P G Q G L K W M G W I N T Y T G E P T Y A D D F K G R F T F T L E T S T S T A Y
Y G M N W V R Q A P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F T I T A E T S T S T L Y
Y G M N W V R Q A P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F T I T A E T S T S T L Y
Y G M N W V R Q A P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F T I T L E T S T S T A Y
Y G M N W V K Q A P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F T I T L E T S T S T A Y
Y G M N W V K Q A P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F A F T L E T S T S T A Y
Y G M N W V R Q A P G K G L K W M G W I N T Y T G E P T Y A D D F K G R F V F S L E T S A S T A F
Y G M N W V X Q X P G X G L X W M G W I N T Y T G E P T Y A D X F X G R X X X X X X T S X S T X X
```

Figure 1C

```
                          CDR3            FR4
                          90           100
- - - - - - - * * * * * * * * - - - - - - - - - -
D A A T Y Y C H Q R S G Y P Y T F G G G T K L E I K
D F A V Y Y C H Q R S G Y P Y T F G G G T K L E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K L E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K L E I K
D X A X Y Y C H Q R S G Y P Y T F G G G T K L E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K V E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K V E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K V E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K V E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K V E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K V E I K
D A A T Y Y C H Q R S G Y P Y T F G G G T K L E I K
D X A X Y Y C H Q R S G Y P Y T F G G G T K X E I K
```

```
                                          CDR3            FR4
                          90           100              110
- - - - - - - - - - - - - - - - - - * * * * * * * - - - - - - - - - - - -
L Q I N N L R N E D M A T Y F C V R F I S K G D Y W G Q G T S V T V S S
L Q L S S L R A E D T A V Y F C V R F I S K G D Y W G Q G T S V T V S S
L Q I N N L R S E D T A T Y F C V R F I S K G D Y W G Q G T S V T V S S
L Q X X X L R X E D X A X Y F C V R F I S K G D Y W G Q G T S V T V S S
L Q L N N L R S E D T A T Y F C V R F I S K G D Y W G T G T T V T V S S
L Q L N N L R S E D M A T Y F C V R F I S K G D Y W G Q G T T V T V S S
L Q L N N L R S E D T A T Y F C V R F M S K G D Y W G Q G T T V T V S S
L Q L N N L R S E D T A T Y F C V R F I S K G D Y W G Q G T T V T V S S
L Q L N N L R S E D T A T Y F C V R F I S K G D Y W G Q G T T V T V S S
L Q L N N L R S E D T A T Y F C V R F I S K G D Y W G Q G T T V T V S S
L Q L N N L R S E D T A T Y F C V R F I S K G D Y W G Q G T T V T V S S
L Q L N N L R S E D T A T Y F C V R F I S K G D Y W G Q G T S V T V S S
L Q X X X L R X E D X A X Y F C V R F X S K G D Y W G X G T X V T V S S
```

SEQUENCE LISTING

<110> Gillies, Stephen
Lo, Kin-Ming
Qian, Xiugi
Lexigen Pharmaceuticals Corp.

<120> Recombinant Tumor Specific Antibody And Use Thereof

<130> LEX-019PC

<150> US 60/288,564
<151> 2001-05-03

<160> 42

<170> PatentIn version 3.0

<210> 1
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS VK mouse

<400> 1

```
Gln Ile Leu Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly
1               5                   10                  15

Glu Lys Val Thr Met Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
            20                  25                  30

Leu Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys Pro Trp Ile Phe
        35                  40                  45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ala Arg Phe Ser Gly Ser
    50                  55                  60

Gly Ser Gly Thr Ser Tyr Ser Leu Ile Ile Ser Ser Met Glu Ala Glu
65                  70                  75                  80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
                85                  90                  95

Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
            100                 105
```

<210> 2
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS VH mouse

<400> 2

```
Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu
1               5                   10                  15
```

```
Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
            20                  25                  30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
        35                  40                  45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
    50                  55                  60

Lys Gly Arg Phe Val Phe Ser Leu Glu Thr Ser Ala Ser Thr Ala Phe
65                  70                  75                  80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
                85                  90                  95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Ser Val
            100                 105                 110

Thr Val Ser Ser
        115


<210>  3
<211>  106
<212>  PRT
<213>  Artificial sequence

<220>
<223>  variable light chain sequence in the EpCAM antibody

<220>
<221>  misc_feature
<222>  (1)..(1)
<223>  wherein Xaa at position 1 is a glutamic acid


<220>
<221>  misc_feature
<222>  (3)..(3)
<223>  wherein Xaa at position 3 is a valine


<220>
<221>  misc_feature
<222>  (10)..(10)
<223>  wherein Xaa at position 10 is a threonine or a serine


<220>
<221>  misc_feature
<222>  (11)..(11)
<223>  wherein Xaa at position 11 is a leucine


<220>
<221>  misc_feature
<222>  (12)..(12)
<223>  wherein Xaa at position 12 is an alanine
```

<220>
<221> misc_feature
<222> (13)..(13)
<223> wherein Xaa at position 13 is a leucine or a valine

<220>
<221> misc_feature
<222> (17)..(17)
<223> wherein Xaa at position 17 is a glutamine

<220>
<221> misc_feature
<222> (18)..(18)
<223> wherein Xaa at position 18 is an arginine

<220>
<221> misc_feature
<222> (19)..(19)
<223> wherein Xaa at position 19 is an alanine

<220>
<221> misc_feature
<222> (21)..(21)
<223> wherein Xaa at position 21 is a leucine or an isoleucine

<220>
<221> misc_feature
<222> (32)..(32)
<223> wherein Xaa at position 32 is an isoleucine

<220>
<221> misc_feature
<222> (36)..(36)
<223> wherein Xaa at position 36 is a leucine

<220>
<221> misc_feature
<222> (41)..(41)
<223> wherein Xaa at position 41 is a glutamine

<220>
<221> misc_feature
<222> (42)..(42)
<223> wherein Xaa at position 42 is an alanine or a proline

<220>
<221> misc_feature
<222> (45)..(45)
<223> wherein Xaa at position 45 is a leucine

<220>
<221> misc_feature

<222> (46)..(46)
<223> wherein Xaa at position 46 is a leucine

<220>
<221> misc_feature
<222> (48)..(48)
<223> wherein Xaa at position 48 is a tyrosine

<220>
<221> misc_feature
<222> (57)..(57)
<223> wherein Xaa at position 57 is an isoleucine

<220>
<221> misc_feature
<222> (59)..(59)
<223> wherein Xaa at position 59 is a serine

<220>
<221> misc_feature
<222> (69)..(69)
<223> wherein Xaa at position 69 is a aspartic acid or a threonine

<220>
<221> misc_feature
<222> (71)..(71)
<223> wherein Xaa at position 71 is a threonine

<220>
<221> misc_feature
<222> (73)..(73)
<223> wherein Xaa at position 73 is a threonine

<220>
<221> misc_feature
<222> (75)..(75)
<223> wherein Xaa at position 75 is an asparagine

<220>
<221> misc_feature
<222> (77)..(77)
<223> wherein Xaa at position 77 is a leucine

<220>
<221> misc_feature
<222> (79)..(79)
<223> wherein Xaa at position 79 is a proline

<220>
<221> misc_feature

<222> (82)..(82)
<223> wherein Xaa at position 82 is a phenylalanine

<220>

<221> misc_feature
<222> (84)..(84)
<223> wherein Xaa at position 84 is a valine

<220>
<221> misc_feature
<222> (103)..(103)
<223> wherein Xaa at position 103 is a valine

<400> 3

Xaa Ile Xaa Leu Thr Gln Ser Pro Ala Xaa Xaa Xaa Xaa Ser Pro Gly
1 5 10 15

Xaa Xaa Xaa Thr Xaa Thr Cys Ser Ala Ser Ser Ser Val Ser Thr Xaa
20 25 30

Leu Trp Tyr Xaa Gln Lys Pro Gly Xaa Xaa Pro Lys Xaa Xaa Ile Xaa
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Xaa Pro Xaa Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Xaa Tyr Xaa Leu Xaa Ile Xaa Ser Xaa Glu Xaa Glu
65 70 75 80

Asp Xaa Ala Xaa Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Xaa Glu Ile Lys
100 105

<210> 4
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> variable heavy chain sequence in the EpCAM antibody

<220>
<221> misc_feature
<222> (2)..(2)
<223> wherein Xaa at position 2 is an isoleucine or a valine

<220>
<221> misc_feature
<222> (9)..(9)
<223> wherein Xaa at position 9 is a proline or an alanine

<220>
<221> misc_feature

<222> (11)..(11)
<223> wherein Xaa at position 11 is a leucine or a valine

<220>

<221> misc_feature
<222> (16)..(16)
<223> wherein Xaa at position 16 is a glutamic acid or a serine

<220>
<221> misc_feature
<222> (17)..(17)
<223> wherein Xaa at position 17 is a threonine or a serine

<220>
<221> misc_feature
<222> (38)..(38)
<223> wherein Xaa at position 38 is a lysine or an arginine

<220>
<221> misc_feature
<222> (40)..(40)
<223> wherein Xaa at position 40 is a threonine or an alanine

<220>
<221> misc_feature
<222> (43)..(43)
<223> wherein Xaa at position 43 is a lysine or a glutamine

<220>
<221> misc_feature
<222> (46)..(46)
<223> wherein Xaa at position 46 is a lysine or a glutamic acid

<220>
<221> msic_feature
<222> (63)..(63)
<223> wherein Xaa at position 63 is an aspartic acid or a lysine

<220>
<221> misc_feature
<222> (65)..(65)
<223> wherein Xaa at position 65 is a lysine or a glutamine

<220>
<221> misc_feature
<222> (68)..(68)
<223> wherein Xaa at position 68 is a phenylalanine or a valine

<220>
<221> misc_feature
<222> (69)..(69)

<223> wherein Xaa at position 69 is an alanine, a threonine or a valine

<220>

<221> misc_feature
<222> (70)..(70)
<223> wherein Xaa at position 70 is a phenylalanine or an isoleucine

<220>
<221> misc_feature
<222> (71)..(71)
<223> wherein Xaa at position 71 is a serine or a threonine

<220>
<221> misc_feature
<222> (72)..(72)
<223> wherein Xaa at position 72 is a leucine or an alanine

<220>
<221> misc_feature
<222> (73)..(73)
<223> wherein Xaa at position 73 is a glutamic acid or an aspartic acid

<220>
<221> misc_feature
<222> (76)..(76)
<223> wherein Xaa at position 76 is an alanine or a threonine

<220>
<221> misc_feature
<222> (79)..(79)
<223> wherein Xaa at position 79 is an alanine or a leucine

<220>
<221> misc_feature
<222> (80)..(80)
<223> wherein Xaa at position 80 is a phenylalanine or a tyrosine

<220>
<221> misc_feature
<222> (83)..(83)
<223> wherein Xaa at position 83 is an isoleucine of a leucine

<220>
<221> misc_feature
<222> (84)..(84)
<223> wherein Xaa at position 84 is an asparagine or a serine

<220>
<221> misc_feature

<222> (85)..(85)
<223> wherein Xaa at position 85 is an asparagine or a serine

<220>
<221> misc_feature
<222> (88)..(88)
<223> wherein Xaa at position 88 is an asparagine, an alanine or a serine

<220>
<221> misc_feature
<222> (91)..(91)
<223> wherein Xaa at position 91 is a methionine or a threonine

<220>
<221> misc_feature
<222> (93)..(93)
<223> wherein Xaa at position 93 is a threonine or a valine

<220>
<221> misc_feature
<222> (100)..(100)
<223> wherein Xaa at position 100 is an isoleucine or a methionine

<220>
<221> misc_feature
<222> (108)..(108)
<223> wherein Xaa at position 108 is a glutamine or a threonine

<220>
<221> misc_feature
<222> (111)..(111)
<223> wherein Xaa at position 111 is a serine or a threonine

<400> 4

```
Gln Xaa Gln Leu Val Gln Ser Gly Xaa Glu Xaa Lys Lys Pro Gly Xaa
1               5                   10                  15

Xaa Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
            20                  25                  30

Gly Met Asn Trp Val Xaa Gln Xaa Pro Gly Xaa Gly Leu Xaa Trp Met
        35                  40                  45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Xaa Phe
    50                  55                  60

Xaa Gly Arg Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser Xaa Ser Thr Xaa Xaa
65                  70                  75                  80

Leu Gln Xaa Xaa Xaa Leu Arg Xaa Glu Asp Xaa Ala Xaa Tyr Phe Cys
                85                  90                  95
```

Val Arg Phe Xaa Ser Lys Gly Asp Tyr Trp Gly Xaa Gly Thr Xaa Val
100 105 110

Thr Val Ser Ser

115

<210> 5
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> light sequence consensus
<220>
<221> misc_feature

<222> (1)..(1)
<223> wherein xaa at position 1 is a glutamine or a glutamic acid

<220>
<221> misc_feature
<222> (3)..(3)
<223> wherein Xaa at position 3 is a leucine or a valine

<220>
<221> misc_feature
<222> (10)..(10)
<223> wherein Xaa at position 10 is an isoleucine or a threonine

<220>
<221> misc_feature
<222> (11)..(11)
<223> wherein Xaa at position 11 is a methionine or a leucine

<220>
<221> misc_feature
<222> (13)..(13)
<223> wherein Xaa at position 13 is an alanine or a leucine

<220>
<221> misc_feature
<222> (18)..(18)
<223> wherein Xaa at position 18 is a lysine or an arginine

<220>
<221> misc_feature
<222> (21)..(21)
<223> wherein Xaa at position 21 is a methionine or a leucine

<220>
<221> misc_feature
<222> (41)..(41)
<223> wherein Xaa at position 41 is a serine or a glutamine

<220>
<221> misc_feature
<222> (42)..(42)

<223> wherein Xaa at position 42 is a serine or an alanine

<220>
<221> misc_feature
<222> (45)..(45)
<223> wherein Xaa at position 45 is a proline or a leucine

<220>
<221> misc_feature
<222> (46)..(46)
<223> wherein Xaa at position 46 is a tryptophan or a leucine

<220>
<221> misc_feature
<222> (57)..(57)
<223> wherein Xaa at position 57 is a phenylalanine or an isoleucine

<220>
<221> misc_feature
<222> (69)..(69)
<223> wherein Xaa at position 69 is a serine or an aspartic acid

<220>
<221> misc_feature
<222> (71)..(71)
<223> wherein Xaa at position 71 is a serine or a threonine

<220>
<221> misc_feature
<222> (73)..(73)
<223> wherein Xaa at position 73 is an isoleucine or a threonine

<220>
<221> misc_feature
<222> (77)..(77)
<223> wherein Xaa at position 77 is a methionine or a leucine

<220>
<221> misc_feature
<222> (79)..(79)
<223> wherein Xaa at position 79 is an alanine or a proline

<220>
<221> misc_feature
<222> (82)..(82)
<223> wherein Xaa at position 82 is an alanine or a phenylalanine

<220>
<221> misc_feature
<222> (84)..(84)
<223> wherein Xaa at position 84 is a threonine or a valine

<400> 5

```
Xaa Ile Xaa Leu Thr Gln Ser Pro Ala Xaa Xaa Ser Xaa Ser Pro Gly
1               5                   10                  15

Glu Xaa Val Thr Xaa Thr Cys Ser Ala Ser Ser Ser Val Ser Thr Met
            20                  25                  30

Leu Trp Tyr Gln Gln Lys Pro Gly Xaa Xaa Pro Lys Xaa Xaa Ile Phe
        35                  40                  45

Asp Thr Ser Asn Leu Ala Ser Gly Xaa Pro Ala Arg Phe Ser Gly Ser
    50                  55                  60

Gly Ser Gly Thr Xaa Tyr Xaa Leu Xaa Ile Ser Ser Xaa Glu Xaa Glu
65                  70                  75                  80

Asp Xaa Ala Xaa Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
                85                  90                  95

Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
            100                 105
```

<210> 6
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> heavy sequence consensus

<220>
<221> misc_feature
<222> (2)..(2)
<223> wherein Xaa at position 2 is an isoleucine or a valine

<220>
<221> misc_feature
<222> (9)..(9)
<223> wherein Xaa at position 9 is a proline or an alanine

<220>
<221> misc_feature
<222> (11)..(11)
<223> wherein Xaa at position 11 is a leucine or a valine

<220>
<221> misc_feature
<222> (17)..(17)
<223> wherein Xaa at position 17 is a threonine or a serine

<220>
<221> misc_feature
<222> (38)..(38)
<223> wherein Xaa at position 38 is a lysine or an arginine

<220>
<221> misc_feature
<222> (40)..(40)
<223> wherein Xaa at position 40 is a threonine or an alanine

<220>
<221> misc_feature
<222> (46)..(46)
<223> wherein Xaa at position 46 is a lysine or a glutamic acid

<220>
<221> misc_feature
<222> (63)..(63)
<223> wherein Xaa at position 63 is an aspartic acid or a lysine

<220>
<221> misc_feature
<222> (65)..(65)
<223> wherein Xaa at position 65 is a lysine or a glutamine

<220>
<221> misc_feature
<222> (68)..(68)
<223> wherein Xaa at position 68 is a phenylalanine or a valine

<220>
<221> misc_feature
<222> (69)..(69)
<223> wherein Xaa at position 69 is an alanine or a threonine

<220>
<221> misc_feature
<222> (70)..(70)
<223> wherein Xaa at position 70 is a phenylalanine or an isoleucine

<220>
<221> misc_feature
<222> (73)..(73)
<223> wherein Xaa at position 73 is a glutamic acid or an aspartic acid

<220>
<221> misc_feature
<222> (76)..(76)
<223> wherein Xaa at position 76 is an alanine or a threonine

<220>
<221> misc_feature
<222> (80)..(80)

<223> wherein Xaa at position 80 is a phenylalanine or a tyrosine

<220>
<221> misc_feature
<222> (83)..(83)
<223> wherein Xaa at position 83 is an isoleucine or a leucine

<220>
<221> misc_feature
<222> (84)..(84)
<223> wherein Xaa at position 84 is an asparagine or a serine

<220>
<221> misc_feature
<222> (85)..(85)
<223> wherein Xaa at position 85 is an asparagine or a serine

<220>
<221> misc_feature
<222> (88)..(88)
<223> wherein Xaa at position 88 is an asparagine, an alanine or a serine

<220>
<221> misc_feature
<222> (91)..(91)
<223> wherein Xaa at position 91 is a methionine or a threonine

<220>
<221> misc_feature
<222> (93)..(93)
<223> wherein Xaa at position 93 is a threonine or a valine

<220>
<221> misc_feature
<222> (108)..(108)
<223> wherein Xaa at position 108 is a glutamine or a threonine

<400> 6

```
Gln Xaa Gln Leu Val Gln Ser Gly Xaa Glu Xaa Lys Lys Pro Gly Glu
1               5                   10                  15

Xaa Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
            20                  25                  30

Gly Met Asn Trp Val Xaa Gln Xaa Pro Gly Lys Gly Leu Xaa Trp Met
        35                  40                  45
```

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Xaa Phe
50 55 60

Xaa Gly Arg Xaa Xaa Xaa Ser Leu Xaa Thr Ser Xaa Ser Thr Ala Xaa
65 70 75 80

Leu Gln Xaa Xaa Xaa Leu Arg Xaa Glu Asp Xaa Ala Xaa Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Xaa Gly Thr Ser Val
100 105 110

Thr Val Ser Ser
115

<210> 7
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> Vk6 light chain

<400> 7

Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
1 5 10 15

Glu Arg Val Thr Leu Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

Leu Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Lys Leu Leu Ile Phe
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Ile Pro Ala Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile Ser Ser Leu Glu Pro Glu
65 70 75 80

Asp Phe Ala Val Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
100 105

<210> 8
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> VK7 light chain

<400> 8
Gln Ile Leu Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly
1 5 10 15

Glu Arg Val Thr Met Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

Leu Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys Pro Trp Ile Phe
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ala Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Ser Tyr Ser Leu Ile Ile Ser Ser Met Glu Pro Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
100 105

<210> 9
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> VK8 light chain

<400> 9

Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
1 5 10 15

Glu Arg Val Thr Leu Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

Leu Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys Pro Trp Ile Phe
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ala Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Ser Tyr Ser Leu Ile Ile Ser Ser Met Glu Ala Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
100 105

<210> 10
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS VK veneered

<400> 10

Gln Ile Leu Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Pro Gly
1 5 10 15

Gln Arg Ala Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

Leu Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Pro Trp Ile Phe
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ala Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Ser Tyr Thr Leu Thr Ile Asn Ser Leu Glu Ala Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Val Glu Ile Lys
100 105

<210> 11
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VK1

<400> 11

Gln Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Pro Gly
1 5 10 15

Gln Arg Ala Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Ile
20 25 30

Leu Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Pro Trp Ile Phe
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ser Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Ser Tyr Thr Leu Thr Ile Asn Ser Leu Glu Ala Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Val Glu Ile Lys
100 105

<210> 12
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VK2

<400> 12

```
Gln Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Pro Gly
1               5                   10                  15

Gln Arg Ala Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
            20                  25                  30

Leu Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Pro Trp Ile Phe
        35                  40                  45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ser Arg Phe Ser Gly Ser
    50                  55                  60

Gly Ser Gly Thr Ser Tyr Thr Leu Thr Ile Asn Ser Leu Glu Ala Glu
65                  70                  75                  80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
                85                  90                  95

Phe Gly Gly Gly Thr Lys Val Glu Ile Lys
            100                 105
```

<210> 13
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS-deimmunized VK3

<400> 13

```
Gln Ile Leu Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Pro Gly
1               5                   10                  15

Gln Arg Ala Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
            20                  25                  30

Leu Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Pro Trp Ile Phe
        35                  40                  45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ser Arg Phe Ser Gly Ser
    50                  55                  60

Gly Ser Gly Thr Ser Tyr Thr Leu Thr Ile Asn Ser Leu Glu Ala Glu
65                  70                  75                  80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
                85                  90                  95

Phe Gly Gly Gly Thr Lys Val Glu Ile Lys
            100                 105
```

<210> 14
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VK4

<400> 14

Gln Ile Leu Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Pro Gly
1 5 10 15

Gln Arg Ala Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

Leu Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Pro Trp Ile Phe
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ala Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Ser Tyr Thr Leu Thr Ile Asn Ser Leu Glu Ala Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Val Glu Ile Lys
100 105

<210> 15
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VK5

<400> 15

Gln Ile Leu Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Pro Gly
1 5 10 15

Gln Arg Ala Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

Leu Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys Pro Trp Ile Tyr
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ala Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Ser Tyr Thr Leu Thr Ile Asn Ser Leu Glu Ala Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Val Glu Ile Lys
100 105

<210> 16
<211> 106
<212> PRT
<213> Artificial sequence

<220>
<223> KS VK mouse

<400> 16

Gln Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Ala Ser Pro Gly
1 5 10 15

Glu Arg Val Thr Ile Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

Leu Trp Tyr Leu Gln Lys Pro Gly Ser Ser Pro Lys Pro Trp Ile Phe
35 40 45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ser Arg Phe Ser Gly Ser
50 55 60

Gly Ser Gly Thr Thr Tyr Ser Leu Ile Ile Ser Ser Leu Glu Ala Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
85 90 95

Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
100 105

<210> 17
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> VH6 heavy chain

<400> 17

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Glu
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Gln Lys Phe
50 55 60

Gln Gly Arg Val Thr Ile Ser Leu Asp Thr Ser Thr Ser Thr Ala Tyr
65 70 75 80

Leu Gln Leu Ser Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Ser Val
100 105 110

Thr Val Ser Ser
115

<210> 18
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> VH7 heavy chain

<400> 18
Gln Ile Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Glu
1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Lys Gln Thr Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Glu Thr Ser Thr Ser Thr Ala Phe
65 70 75 80

Leu Gln Ile Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Ser Val
100 105 110

Thr Val Ser Ser
115

<210> 19
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> VH2.5 heavy chain

<400> 19

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Thr Ile Thr Ala Glu Thr Ser Thr Ser Thr Leu Tyr
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Thr Gly Thr Thr Val
100 105 110

Thr Val Ser Ser

115

<210> 20
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> KS VH veneered

<400> 20

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Lys Gln Ala Pro Gly Gln Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Thr Phe Thr Ile Glu Thr Ser Thr Ser Thr Ala Tyr
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Met Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser
115

<210> 21
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VH1

<400> 21

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Thr Ile Thr Ala Glu Thr Ser Thr Ser Thr Leu Tyr
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Met Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser
115

<210> 22
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VH2

<400> 22

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Thr Ile Thr Ala Glu Thr Ser Thr Ser Thr Leu Tyr
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser
115

<210> 23
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VH3

<400> 23

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Thr Ile Thr Leu Glu Thr Ser Thr Ser Thr Ala Tyr
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser
115

<210> 24
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> KS- deimmunized VH4

<400> 24

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Ser
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Thr Ile Thr Leu Glu Thr Ser Thr Ser Thr Ala Tyr
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser
115

<210> 25
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> KS de-immunized VH5

<400> 25

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Lys Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Thr Leu Glu Thr Ser Thr Ser Thr Ala Tyr
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Thr Val
100 105 110

Thr Val Ser Ser
115

<210> 26
<211> 116
<212> PRT
<213> Artificial sequence

<220>
<223> KS VH mouse

<400> 26

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu
1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Val Phe Ser Leu Glu Thr Ser Ala Ser Thr Ala Phe
65 70 75 80

Leu Gln Leu Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Ser Val
100 105 110

Thr Val Ser Ser
115

<210> 27
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> KSA sense primer

<400> 27
tctagagcag catggcgccc ccgca 25

<210> 28
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> KSA antisense primer

<400> 28
ctcgagttat gcattgagtt ccct 24

<210> 29
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> linker-adapter

<400> 29
aattctcaat gcagggc 17

<210> 30
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> linker-adapter

<400> 30
gagttacgtc ccgaatt 17

<210> 31
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> VH forward primer

<400> 31
gactcgagcc caagtcttag acatc 25

<210> 32
<211> 33
<212> DNA

<213> Artificial sequence

<220>
<223> VH reverse primer

<400> 32
caagcttacc tgaggagacg gtgactgacg ttc 33

<210> 33
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> VK forward primer

<400> 33
gatctagaca agatggattt tcaagtg 27

<210> 34
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> VK reverse primer

<400> 34
gaagatctta cgttttattt ccagcttgg 29

<210> 35
<211> 117
<212> PRT
<213> Artificial sequence

<220>
<223> VH369 heavy chain

<400> 35

Gln Ile Gln Leu Val Gln Ser Gly Pro Glu Leu Lys Lys Pro Gly Glu
1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Lys Gln Thr Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Glu Thr Ser Ala Ser Thr Ala Phe
65 70 75 80

Leu Gln Ile Gln Gln Pro Gln Asn Met Arg Thr Met Ala Thr Tyr Phe
85 90 95

Cys Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Ser
100 105 110

Val Thr Val Ser Ser
115

<210> 36
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> VH2.1 partial sequence

<400> 36

Ala Thr Tyr Phe Cys Val Arg Phe Ile Ile Ser Lys Gly Asp Tyr Trp
1 5 10 15

Gly Gln Gly

<210> 37
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> VH2.2 partial sequence

<400> 37

Ala Thr Tyr Phe Cys Val Arg Phe Ile Val Ser Lys Gly Asp Tyr Trp
1 5 10 15

Gly Gln Gly

<210> 38
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> VH2.3 partial sequence

<400> 38

Ala Thr Tyr Phe Cys Val Arg Phe Ile Ser Ala Lys Gly Asp Tyr Trp
1 5 10 15

Gly Gln Gly

<210> 39
<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> VH2.4 partial sequence

<400> 39

Ala Thr Tyr Phe Cys Val Arg Phe Ile Ser Lys Thr Gly Asp Tyr Trp
1   5   10   15

Gly Gln Gly

<210> 40
<211> 10494
<212> DNA
<213> Artificial sequence

<220>
<223> pdHL7s-VK8/VH7-IL2 sequence

<400> 40
gtcgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 60

gtcgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 120

gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc 180

ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 240

ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 300

atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg 360

cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg 420

tattagtcat cgctattacc atggtgatgc ggttttggca gtacatcaat gggcgtggat 480

agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 540

tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc 600

aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctctc tggctaacta 660

cagaacccac tgcttactgg cttatcgaaa ttaatacgac tcactatagg gagaccctct 720

agaatgaagt tgcctgttag gctgttggtg ctgatgttct ggattcctgg tgaggagaga 780

gggaagtgag ggaggagaat ggacagggag caggagcact gaatcccatt gctcattcca 840

tgtatctggc atgggtgaga agatgggtct tatcctccag catggggcct ctggggtgaa 900

tacttgttag agggaggttc cagatgggaa catgtgctat aatgaagatt atgaaatgga 960

tgcctgggat ggtctaagta atgccttaga agtgactaga cacttgcaat tcactttttt 1020

tggtaagaag agatttttag gctataaaaa aatgttatgt aaaaataaac gatcacagtt 1080

gaaataaaaa aaaaatataa ggatgttcat gaattttgtg tataactatg tatttctctc 1140

tcattgtttc agcttcctta agcgagatcg tgctgaccca gtcccccgcc accctgtccc 1200

tgtcccccgg cgagcgcgtg accctgacct gctccgcctc ctcctccgtg tcctacatgc 1260

tgtggtacca gcagaagcca ggatcctcgc ccaaaccctg gatttttgac acatccaacc 1320

tggcttctgg attccctgct cgcttcagtg gcagtgggtc tgggacctct tactctctca 1380
taatcagcag catggaggct gaagatgctg ccacttatta ctgccatcag cggagtggtt 1440
acccgtacac gttcggaggg gggaccaagc tggaaataaa acgtaagatc ccgcaattct 1500
aaactctgag gggtcggat gacgtggcca ttctttgcct aaagcattga gtttactgca 1560
aggtcagaaa agcatgcaaa gccctcagaa tggctgcaaa gagctccaac aaaacaattt 1620
agaactttat taaggaatag ggggaagcta ggaagaaact caaaacatca agattttaaa 1680
tacgcttctt ggtctccttg ctataattat ctgggataag catgctgttt tctgtctgtc 1740
cctaacatgc cctgtgatta tccgcaaaca acacacccaa gggcagaact ttgttactta 1800
aacaccatcc tgtttgcttc tttcctcagg aactgtggct gcaccatctg tcttcatctt 1860
cccgccatct gatgagcagt tgaaatctgg aactgcctct gttgtgtgcc tgctgaataa 1920
cttctatccc agagaggcca aagtacagtg gaaggtggat aacgccctcc aatcgggtaa 1980
ctcccaggag agtgtcacag agcaggacag caaggacagc acctacagcc tcagcagcac 2040
cctgacgctg agcaaagcag actacgagaa acacaaagtc tacgcctgcg aagtcaccca 2100
tcagggcctg agctcgcccg tcacaaagag cttcaacagg ggagagtgtt agagggagaa 2160
gtgcccccac ctgctcctca gttccagcct gacccctcc catcctttgg cctctgaccc 2220
tttttccaca ggggacctac ccctattgcg gtcctccagc tcatctttca cctcaccccc 2280
ctcctcctcc ttggctttaa ttatgctaat gttggaggag aatgaataaa taaagtgaat 2340
ctttgcacct gtggtttctc tctttcctca atttaataat tattatctgt tgtttaccaa 2400
ctactcaatt tctcttataa gggactaaat atgtagtcat cctaaggcgc ataaccattt 2460
ataaaaatca tccttcattc tattttaccc tatcatcctc tgcaagacag tcctccctca 2520
aacccacaag ccttctgtcc tcacagtccc ctgggccatg gtaggagaga cttgcttcct 2580
tgttttcccc tcctcagcaa gccctcatag tcctttttaa gggtgacagg tcttacggtc 2640
atatatcctt tgattcaatt ccctgggaat caaccaaggc aaatttttca aaagaagaaa 2700
cctgctataa agagaatcat tcattgcaac atgatataaa ataacaacac aataaaagca 2760
attaaataaa caaacaatag ggaaatgttt aagttcatca tggtacttag acttaatgga 2820
atgtcatgcc ttatttacat ttttaaacag gtactgaggg actcctgtct gccaagggcc 2880
gtattgagta ctttccacaa cctaatttaa tccacactat actgtgagat taaaaacatt 2940
cattaaaatg ttgcaaaggt tctataaagc tgagagacaa atatattcta taactcagca 3000
atcccacttc tagggtcgac gttgacattg attattgact agttattaat agtaatcaat 3060
tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa 3120

tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 3180

tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 3240

aactgcccac ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt 3300

caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc 3360

tacttggcag tacatctacg tattagtcat cgctattacc atggtgatgc ggttttggca 3420

gtacatcaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat 3480

tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa 3540

caactccgcc ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag 3600

cagagctctc tggctaacta cagaacccac tgcttactgg cttatcgaaa ttaatacgac 3660

tcactatagg gagacccaag ctcctcgagg ctagaatgaa gttgcctgtt aggctgttgg 3720

tgctgatgtt ctggattcct ggtgaggaga gagggaagtg agggaggaga atggacaggg 3780

agcaggagca ctgaatccca ttgctcattc catgtatctg gcatgggtga gaagatgggt 3840

cttatcctcc agcatggggc ctctggggtg aatacttgtt agagggaggt tccagatggg 3900

aacatgtgct ataatgaaga ttatgaaatg gatgcctggg atggtctaag taatgcctta 3960

gaagtgacta gacacttgca attcactttt tttggtaaga agagattttt aggctataaa 4020

aaaatgttat gtaaaaataa acgatcacag ttgaaataaa aaaaaaatat aaggatgttc 4080

atgaattttg tgtataacta tgtatttctc tctcattgtt tcagcttcct taagccagat 4140

ccagttggtg cagtctggag ctgaggtgaa gaagcctgga gagacagtca agatctcctg 4200

caaggcttct gggtatacct tcacaaacta tggaatgaac tgggtgaagc agactccagg 4260

aaagggttta aagtggatgg gctggataaa cacctacact ggagaaccaa catatgctga 4320

tgacttcaag ggacggtttg ccttctcttt ggaaacctct accagcactg cctttttgca 4380

gatcaacaat ctcagaagtg aggacacggc tacatatttc tgtgtaagat ttatttctaa 4440

gggggactac tggggtcaag gaacgtcagt caccgtctcc tcaggtaagc tttctggggc 4500

aggccaggcc tgaccttggc tttggggcag ggagggggct aaggtgaggc aggtggcgcc 4560

agccaggtgc acacccaatg cccatgagcc cagacactgg acgctgaacc tcgcggacag 4620

ttaagaaccc aggggcctct gcgccctggg cccagctctg tcccacaccg cggtcacatg 4680

gcaccacctc tcttgcagcc tccaccaagg gcccatcggt cttccccctg gcaccctcct 4740

ccaagagcac ctctgggggc acagcggccc tgggctgcct ggtcaaggac tacttccccg 4800

aaccggtgac ggtgtcgtgg aactcaggcg ccctgaccag cggcgtgcac accttcccgg 4860

ctgtcctaca gtcctcagga ctctactccc tcagcagcgt ggtgaccgtg ccctccagca 4920

gcttgggcac ccagacctac atctgcaacg tgaatcacaa gcccagcaac accaaggtgg 4980

acaagagagt tggtgagagg ccagcacagg gagggagggt gtctgctgga agccaggctc 5040
agcgctcctg cctggacgca tcccggctat gcagtcccag tccagggcag caaggcaggc 5100
cccgtctgcc tcttcacccg gaggcctctg cccgccccac tcatgctcag ggagagggtc 5160
ttctggcttt ttccccaggc tctgggcagg cacaggctag gtgcccctaa cccaggccct 5220
gcacacaaag gggcaggtgc tgggctcaga cctgccaaga gccatatccg ggaggaccct 5280
gcccctgacc taagcccacc ccaaaggcca aactctccac tccctcagct cggacacctt 5340
ctctcctccc agattccagt aactcccaat cttctctctg cagagcccaa atcttgtgac 5400
aaaactcaca catgcccacc gtgcccaggt aagccagccc aggcctcgcc ctccagctca 5460
aggcgggaca ggtgccctag agtagcctgc atccagggac aggccccagc cgggtgctga 5520
cacgtccacc tccatctctt cctcagcacc tgaactcctg gggggaccgt cagtcttcct 5580
cttcccccca aaacccaagg acaccctcat gatctcccgg acccctgagg tcacatgcgt 5640
ggtggtggac gtgagccacg aagaccctga ggtcaagttc aactggtacg tggacggcgt 5700
ggaggtgcat aatgccaaga caaagccgcg ggaggagcag tacaacagca cgtaccgtgt 5760
ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat ggcaaggagt acaagtgcaa 5820
ggtctccaac aaagccctcc cagcccccat cgagaaaacc atctccaaag ccaaaggtgg 5880
gacccgtggg gtgcgagggc cacatggaca gaggccggct cggcccaccc tctgccctga 5940
gagtgaccgc tgtaccaacc tctgtcccta cagggcagcc ccgagaacca caggtgtaca 6000
ccctgccccc atcacgggag gagatgacca agaaccaggt cagcctgacc tgcctggtca 6060
aaggcttcta tcccagcgac atcgccgtgg agtgggagag caatgggcag ccggagaaca 6120
actacaagac cacgcctccc gtgctggact ccgacggctc cttcttcctc tatagcaagc 6180
tcaccgtgga caagagcagg tggcagcagg ggaacgtctt ctcatgctcc gtgatgcatg 6240
aggctctgca caaccactac acgcagaaga gcctctccct gtccccgggt aaagccccaa 6300
cttcaagttc tacaaagaaa acacagctgc aactggagca tctcctgctg gatctccaga 6360
tgattctgaa tggaattaac aactacaaga atcccaaact caccaggatg ctcacattca 6420
agttctacat gcccaagaag gccacagagc tcaaacatct ccagtgtcta gaggaggaac 6480
tcaaacctct ggaggaagtg ctaaacctcg ctcagagcaa aaacttccac ttaagaccta 6540
gggacttaat cagcaatatc aacgtaatag ttctggaact aaagggatcc gaaacaacat 6600
tcatgtgtga atatgctgat gagacagcaa ccattgtaga attcctaaac agatggatta 6660
ccttttgtca aagcatcatc tcaacactaa cttgataatt aagtgctcga gggatccaga 6720
catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg 6780

ctttatttgt gaaatttgtg atgctattgc tttatttgta accattagaa gctgcaataa 6840

acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtggga 6900

ggttttttaa agcaagtaaa acctctacaa atgtggtatg gctgattatg atcctgcctc 6960

gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc agctcccgga gacggtcaca 7020

gcttgtctgt aagcggatgc cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt 7080

ggcgggtgtc ggggcgcagc catgacccag tcacgtagcg atagcggagt gtatactggc 7140

ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac 7200

cgcacagatg cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc tcgctcactg 7260

actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa 7320

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 7380

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 7440

ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 7500

aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 7560

cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcaatgct 7620

cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 7680

aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 7740

cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 7800

ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 7860

ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta 7920

gctcttgatc cggcaaacaa accaccgctg gtagcggtgg tttttttgtt tgcaagcagc 7980

agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 8040

acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 8100

tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 8160

agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 8220

gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg 8280

agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc 8340

cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 8400

ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 8460

cagttaatag tttgcgcaac gttgttgcca ttgctgcagg catcgtggtg tcacgctcgt 8520

cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc 8580

ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt 8640

tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 8700
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 8760
gtatgcggcg accgagttgc tcttgcccgg cgtcaacacg ggataatacc gcgccacata 8820
gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 8880
tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag 8940
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 9000
aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt 9060
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 9120
aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag 9180
aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 9240
ttcaagaatt ccgatccaga catgataaga tacattgatg agtttggaca aaccacaact 9300
agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc tttatttgta 9360
accattagaa gctgcaataa acaagttaac aacaacaatt gcattcattt tatgtttcag 9420
gttcaggggg aggtgtggga ggttttttaa agcaagtaaa acctctacaa atgtggtatg 9480
gctgattatg atctaaagcc agcaaaagtc ccatggtctt ataaaaatgc atagctttcg 9540
gaggggagca gagaacttga aagcatcttc ctgttagtct ttcttctcgt agaccttaaa 9600
ttcatacttg attccttttt cctcctggac ctcagagagg acgcctgggt attctgggag 9660
aagtttatat ttccccaaat caatttctgg gaaaaacgtg tcactttcaa attcctgcat 9720
gatccttgtc acaaagagtc tgaggtggcc tggttgattc atggcttcct ggtaaacaga 9780
actgcctccg actatccaaa ccatgtctac tttacttgcc aattccggtt gttcaataag 9840
tcttaaggca tcatccaaac ttttggcaag aaaatgagct cctcgtggtg gttctttgag 9900
ttctctactg agaactatat taattctgtc ctttaaaggt cgattcttct caggaatgga 9960
gaaccaggtt ttcctaccca taatcaccag attctgttta ccttccactg aagaggttgt 10020
ggtcattctt tggaagtact tgaactcgtt cctgagcgga ggccagggtc ggtctccgtt 10080
cttgccaatc cccatatttt gggacacggc gacgatgcag ttcaatggtc gaaccatgag 10140
ggcaccaagc tagctttttg caaaagccta ggcctccaaa aaagcctcct cactacttct 10200
ggaatagctc agaggccgag gcggcctcgg cctctgcata aataaaaaaa attagtcagc 10260
catggggcgg agaatgggcg gaactgggcg gagttagggg cgggatgggc ggagttaggg 10320
gcgggactat ggttgctgac taattgagat gcatgctttg catacttctg cctgctgggg 10380
agcctgggga ctttccacac ctggttgctg actaattgag atgcatgctt tgcatacttc 10440

tgcctgctgg ggagcctggg gactttccac accctaactg acacacattc caca 10494

<210> 41
<211> 579

<212> PRT
<213> Artificial sequence

<220>
<223> heavy chain-IL2

<400> 41

Gln Ile Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Glu
1 5 10 15

Thr Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30

Gly Met Asn Trp Val Lys Gln Thr Pro Gly Lys Gly Leu Lys Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Asp Asp Phe
50 55 60

Lys Gly Arg Phe Ala Phe Ser Leu Glu Thr Ser Thr Ser Thr Ala Phe
65 70 75 80

Leu Gln Ile Asn Asn Leu Arg Ser Glu Asp Thr Ala Thr Tyr Phe Cys
85 90 95

Val Arg Phe Ile Ser Lys Gly Asp Tyr Trp Gly Gln Gly Thr Ser Val
100 105 110

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala
115 120 125

Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu
130 135 140

Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly
145 150 155 160

Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser
165 170 175

Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu
180 185 190

Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr
195 200 205

Lys Val Asp Lys Arg Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr
210 215 220

Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
225 230 235 240

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro
245 250 255

Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
260 265 270

Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr
275 280 285

Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val
290 295 300

Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
305 310 315 320

Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
325 330 335

Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
340 345 350

Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
355 360 365

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
370 375 380

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
385 390 395 400

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
405 410 415

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
420 425 430

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Ala Pro
435 440 445

Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu Gln Leu Glu His Leu Leu
450 455 460

Leu Asp Leu Gln Met Ile Leu Asn Gly Ile Asn Asn Tyr Lys Asn Pro
465 470 475 480

Lys Leu Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro Lys Lys Ala
485 490 495

Thr Glu Leu Lys His Leu Gln Cys Leu Glu Glu Glu Leu Lys Pro Leu
500 505 510

Glu Glu Val Leu Asn Leu Ala Gln Ser Lys Asn Phe His Leu Arg Pro
515 520 525

Arg Asp Leu Ile Ser Asn Ile Asn Val Ile Val Leu Glu Leu Lys Gly
530 535 540

Ser Glu Thr Thr Phe Met Cys Glu Tyr Ala Asp Glu Thr Ala Thr Ile
545 550 555 560

Val Glu Phe Leu Asn Arg Trp Ile Thr Phe Cys Gln Ser Ile Ile Ser
565 570 575

Thr Leu Thr

<210> 42
<211> 213

<212> PRT
<213> Artificial sequence

<220>
<223> light chain

<400> 42

```
Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
1               5                   10                  15

Glu Arg Val Thr Leu Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
            20                  25                  30

Leu Trp Tyr Gln Gln Lys Pro Gly Ser Ser Pro Lys Pro Trp Ile Phe
        35                  40                  45

Asp Thr Ser Asn Leu Ala Ser Gly Phe Pro Ala Arg Phe Ser Gly Ser
    50                  55                  60

Gly Ser Gly Thr Ser Tyr Ser Leu Ile Ile Ser Ser Met Glu Ala Glu
65                  70                  75                  80

Asp Ala Ala Thr Tyr Tyr Cys His Gln Arg Ser Gly Tyr Pro Tyr Thr
                85                  90                  95

Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg Thr Val Ala Ala Pro
            100                 105                 110

Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr
        115                 120                 125

Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys
    130                 135                 140

Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu
145                 150                 155                 160

Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser
                165                 170                 175

Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala
            180                 185                 190

Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe
        195                 200                 205

Asn Arg Gly Glu Cys
    210
```